Vol. 38, Parts 3 and 4 December 1951

BIOMETRIKA

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR

E. S. PEARSON

ASSOCIATE EDITORS

M. G. KENDALL JOHN WISHART

in consultation with

HARALD CRAMÉR R. C. GEARY

J. B. S. HALDANE

ISSUED BY

THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON

PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

[*Issued* 4 *December* 1951]

BIOMETRIKA

1901–1951

Biometrika was founded by Karl Pearson and W. F. R. Weldon in consultation with Francis Galton, and the first number appeared in October 1901. Its origin was due to the Royal Society requesting that in papers submitted for publication mathematics should be kept apart from biological applications, but there is no need to discuss that old disagreement now, and anyone who wants to read about it can refer to Pearson's *Life of Francis Galton* and his memoir of Weldon. The important thing is that *Biometrika* was founded and is still alive after fifty years of activity. As individuals grow older some of their ways of life change, and this is true also of journals. In the case of *Biometrika* there is less biological and anthropological work than in many of the earlier numbers, but there are plenty of papers on the theoretical side of statistics containing the arithmetical work with which the subject should properly be associated. The publication of tables which will be helpful in practical calculation has also continued.

There is another side to all such affairs and it is one which, when it relates to individuals, we are apt to be shy of discussing; it is the financial side. In the early days a guarantee fund was provided by the founders and their friends, and the two founders not only did a great deal of work for the journal but dipped their hands in their pockets from time to time. After Weldon's death in 1906 the practical management of *Biometrika* was in Karl Pearson's hands, and he had a good deal of anxiety about it in the early days of the 1914–18 War as the circulation fell rapidly. Variations of the arrangements with the Cambridge University Press, who have printed the journal ever since it started, were made for reasons of economy, and Pearson and Mrs Weldon gave money to help to keep it going. The journal was never run for profit but solely for scientific ends. Various auxiliary publications had been issued from 1914 onwards, such as the *Tables for Statisticians and Biometricians*, *Tables of Gamma and Beta Functions*, etc. In 1935 a trust was set up by Karl Pearson with not less than five trustees, who were to be responsible for the financial side and for sanctioning the reissue of back numbers that had gone out of print, the issue of auxiliary publications, etc. The trust deed gives the trustees power to appoint a Managing Editor who shall *ex officio* become a trustee, and with his agreement they may appoint assistant editors, but they are not responsible for the editorial work, though, of course, they may be consulted if the Managing Editor wishes to do so.

As has been indicated the editorial work in the early days was done by Pearson and Weldon (with some assistance from Davenport) in consultation with Galton. From 1906 Pearson assumed entire editorial responsibility, though he had some help from W. R. Macdonell and from members of his own staff at University College. In 1924 his son, E. S. Pearson, became assistant editor and is now, as readers of the journal know, the Managing Editor, a task which he first undertook in 1936. He has, of course, had help in various ways, and recently two associate editors have been appointed. Below these notes a list is given of those who have served the journal since its foundation; and those who contribute to it or read it may well be grateful for the help given in the past. It may be said, in conclusion, that it is the wish of everyone now concerned with *Biometrika* that it may continue to be of use in the subjects

with which it has been associated for so long, that new authors may be added to the list of contributors, and that books and tables may continue to be made available for the growing number of people who undertake statistical work in so many and such an increasing number of directions.

W.P.E.

KARL PEARSON	Founder and Editor 1901–1936, Trustee 1935–1936
W. F. R. WELDON	Founder and Editor 1901–1906
FRANCIS GALTON	Founder and Consultant 1901–1911
C. B. DAVENPORT	Assistant Editor 1901–1909
W. P. ELDERTON	Assistant Editor 1906–1909, Trustee 1935–
W. R. MACDONELL	Assistant Editor 1906–1909
RAYMOND PEARL	Assistant Editor 1906–1909
EGON S. PEARSON	Assistant Editor 1924–1936, Trustee 1935–, Managing Editor 1936–
GEOFFREY M. MORANT	Assistant Editor 1935–1937, Consultant 1937–1948
DAVID HERON	Trustee 1935–
J. F. TOCHER	Trustee 1935–1945
MARGARET V. PEARSON	Trustee 1936–1944
HARALD CRAMÉR	Consultant 1937–
MAJOR GREENWOOD	Consultant 1937–1949, Trustee 1946–1949
J. B. S. HALDANE	Consultant 1937–
JOHN WISHART	Consultant 1937–1948, Associate Editor 1948–
R. C. GEARY	Consultant 1939–
M. G. KENDALL	Trustee 1945–1948, Associate Editor 1948–
L. H. C. TIPPETT	Trustee 1948–
A. BRADFORD HILL	Trustee 1950–

THE ESTIMATION OF POPULATION PARAMETERS FROM DATA OBTAINED BY MEANS OF THE CAPTURE-RECAPTURE METHOD

I. THE MAXIMUM LIKELIHOOD EQUATIONS FOR ESTIMATING THE DEATH-RATE

By P. H. LESLIE and DENNIS CHITTY

Bureau of Animal Population, Department of Zoological Field Studies, Oxford

CONTENTS

		PAGE
1.	Introduction	269
2.	The chain of expected numbers	271
3.	The effect of grouping on the amount of information	272
4.	A method of grouping involving no loss of information	274
5.	Solution of the equations for estimating the survival factor	276
6.	The design of the sampling experiment	278
7.	Preliminary analysis of the data	280
8.	Results of the sampling experiment	282
9.	The effect of reducing the size of sample	285
10.	Summary	287
	Appendix I. Numerical examples of the solutions of the equations	288
	II. Corrections to allow for accidental deaths	290
	III. Detailed results of the sampling experiment	292
	References	292

1. Introduction

The aim of the present paper is to derive the maximum-likelihood equations for estimating the death-rate among a population of wild animals from data obtained by the capture-recapture method. This method, which, in the form it is discussed here, consists of taking a series of samples from a constantly changing population, was developed in the first place for the analysis of tsetse-fly populations by Jackson (1939, 1948, and earlier papers), who gives various methods of estimating not only the number of individuals, but also the birth-rate and the death-rate during successive intervals of time. The same procedure has been employed in the field by Dowdeswell, Fisher & Ford (1940) in the case of butterflies, and by Fisher & Ford (1947) in studying a population of the moth *Panaxia dominula* over a number of years. These last authors describe a method of estimating the numbers and the death-rate, when the latter is assumed to have remained constant during the sampling period. In some of the applications of this technique to insect populations, the samples were taken at daily or weekly intervals, and the series might be continued for a relatively lengthy period of time. When the method is applied to populations of small mammals, however, it may not be possible to do this. The alternation of breeding and non-breeding seasons in many cases divides the year roughly into two sections, which it may be more convenient to study separately. Moreover, it may not be possible, for various reasons, to sample the population by means of some live-trapping technique at too frequent intervals. Since we were primarily interested in the application of the capture-recapture method to populations of this type, we were thus led to consider the properties of a sampling chain of a relatively short length, but in which at each sample the proportion of recaptures was not necessarily small. Two methods of assembling the observed data are discussed in this paper, and the comparative accuracy of the resulting estimates is illustrated by the results of a sampling experiment.

From the theoretical point of view, the essentials of the problem may be put briefly as follows. Suppose that a population consisting of an unknown and variable number (N_t) of individuals is sampled at equidistant intervals of time $t = 0, 1, 2, \ldots, T$, and that at each sampling R_t individuals are captured, marked and returned to the population. (We may assume for the present, in order to simplify the theoretical development, that no individuals are either accidently killed by the method of capture employed, or removed permanently from the population.) Among the $R_t (t = 0, 1, 2, \ldots, T)$ individuals captured and examined at time t, let

u_t = the number unmarked; and

s_t = the number found bearing one or more marks, i.e. the total number of recaptures, so that $R_t = u_t + s_t$.

It is to be noted that the 'recaptures' are defined here as the number of individual animals recaptured, and not as the total number of marks observed among them.

Some of these s_t recaptures may have been captured only once, some twice, others up to t times previously, and in each class all possible combinations of the marks $t = 0, 1, 2, 3, \ldots (t-1)$ may be found. Thus, for example, at the fourth sampling taken at $t = 3$ the following seven classes of marks might occur: 0, 1, 2, 01, 02, 12, 012. Let

$r_{012\ldots t}$ = the number of individuals recaptured at time t, bearing the previous marks $0, 1, 2, \ldots, (t-1)$.

The problem then is: from the results of such a series of recaptures to determine the death-rate, given that the sampling of the population is entirely at random and that the method of actually marking or tagging the animals does not affect their subsequent chances of recapture.

Evidently for this problem to be soluble, at any rate in a comparatively straightforward fashion, some simplifying assumptions must be made. Although the population sampled will consist possibly of a number of age classes, for each of which the age-specific death-rates may be different, nevertheless it is convenient, following Fisher & Ford (1947), to assume that over the sampling period 0 to T the death-rate per unit of time remains constant. Let

P = the constant survival factor, so that out of a total number of individuals N_t alive at time t, PN_t are alive at time $t+1$.

(We are supposing that the succession of samples is taken at equidistant intervals of time, so that we may adopt this interval as our unit. If the sampling is at unequal intervals, the definitions must be altered somewhat, but no real difficulty arises.) The death-rate in the population as a whole over each interval of time is then $Q = 1 - P$. If we assume that this death-rate is approximately the same for all age classes in the population, then the expectation of life of young individuals, at the age they enter the population at risk of capture, is

$$e_0 = -1/\log_e P.$$

The estimation of the survival factor P is, of course, a necessary and preliminary step in the solution of the general problem, which also includes the estimation of the total numbers alive in the population at time t. We propose, however, at present to confine our attention to this first step and to reserve the discussion of the methods for estimating the total numbers for a later paper.

It may be shown that the problem of the simultaneous estimation of the $N_t (t = 1, 2, \ldots, T)$ and P reduces to that of first estimating the latter from the recaptures of marked individuals only and then using this value to determine the N_t.

2. THE CHAIN OF EXPECTED NUMBERS

If at the initial sampling R_0 individuals are captured, marked and released, it is easy, given the results of each sampling, to write down the expected number of marked individuals in the population falling into the various classes at a particular time t. In doing so, it is assumed that the constant survival factor P applies equally to all possible subgroups in the population. As an illustration, this chain up to $t = 3$, or the fourth sampling, is given below, where at each time t the first column gives the total number of marked individuals expected in the population as a whole, and the second column represents the number observed in the sample of R_t taken at that time. The method of constructing the chain of expected numbers is clear from the table, each step being determined merely by the results of the previous samplings. This is, of course, to adopt a purely deterministic model of the succession of events:

$t =$	1		2		3	
	PR_0	r_{01}	$P^2R_0 - Pr_{01}$	r_{02}	$P^3R_0 - P^2r_{01} - Pr_{02}$	r_{03}
			Pu_1	r_{12}	$P^2u_1 - Pr_{12}$	r_{13}
			Pr_{01}	r_{012}	Pu_2	r_{23}
					$P^2r_{01} - Pr_{012}$	r_{013}
					Pr_{02}	r_{023}
					Pr_{12}	r_{123}
					Pr_{012}	r_{0123}
Total	PR_0	s_1	$P^2R_0 + Pu_1$	s_2	$P^3R_0 + P^2u_1 + Pu_2$	s_3

Now, provided that the size of each sample is relatively small compared with the population at risk, which will usually be the case in practice, we may assume that we are dealing with a series of multinomial, instead of the hypergeometric distributions which, strictly speaking, should be considered in this form of sampling from a finite population. Then, at time t, there are $n = 2^t - 1$ possible classes of marks. If p_{it} is the expected proportion in the ith class at time t $(i = 1, 2, \ldots, n; \sum_i p_{it} = 1)$ and r_{it} the number observed in the sample, then the likelihood e^L may be taken, apart from a constant factor, as

$$e^L \propto \prod_i (p_{it})^{r_{it}}.$$

Similarly for all the remaining samples; the combined likelihood is

$$e^L \propto \prod_i (p_{i1})^{r_{i1}} \prod_i (p_{i2})^{r_{i2}} \ldots \prod_i (p_{iT})^{r_{iT}},$$

and the relevant part of the log likelihood is therefore

$$L = \sum_t \sum_i r_{it} \log p_{it},$$

from which, by differentiating with respect to P and equating to zero in the usual way, we obtain the equation for the maximum-likelihood estimate of P.

It will be seen by inspection of this table of expected numbers that in writing down the maximum-likelihood equations it is possible by cancelling to reduce the degree of the polynomials in P for the expected proportions by one. Thus for $t = 2$ we have

$$e^L \propto \left(\frac{PR_0 - r_{01}}{PR_0 + u_1}\right)^{r_{02}} \left(\frac{u_1}{PR_0 + u_1}\right)^{r_{12}} \left(\frac{r_{01}}{PR_0 + u_1}\right)^{r_{012}},$$

and

$$L = r_{02} \log (PR_0 - r_{01}) - s_2 \log (PR_0 + u_1).$$

It is also seen that no estimate can be made of P merely from the results of the first two samplings at $t = 0$ and 1. A minimum of three samples is required for any estimate to be made of the survival factor, and it may easily be shown (by supposing that P changes to P' during the last interval) that this estimate refers to the interval of time 0–1. Extending the chain further, it follows that the value of P, estimated from the distributions of recaptures, refers to the span of time between the original and penultimate sampling, and that no estimate can be made of the survival factor during the last interval in the chain.

3. The effect of grouping on the amount of information

In the type of data which commonly arises in practice, when the size of sample may be small relative to the total population at risk at some particular time, the number of recaptures is likely also to be small. A set of frequency distributions of the $r_{012\ldots t}$ for $t = 0, 1, 2, \ldots, T$ would then consist largely of zeros or small numbers in the various subclasses. Moreover, the number of possible classes of marks at time t, $n = 2^t - 1$, becomes very large as the length of the chain increases. It is therefore convenient, in order to simplify the actual computations, to adopt some system of grouping the observations.

This might be done in a number of different ways, and the interesting question then arises of the loss of information involved in adopting a particular method of grouping. Clearly the method of choice would be one which both simplified the equations by reducing the number of subclasses, and at the same time minimized any loss of information.

The method adopted by Jackson (1939, 1948) and by Fisher & Ford (1947) is the following. Essentially it consists in forming the distribution of marks observed among the recaptured animals, each individual being counted as many times as it has marks. Thus, from the observed $r_{012\ldots t}$ at time t we form the sums

$$\begin{aligned} \sum_{0} r_{012\ldots t} &= k_{0t}, \text{ say,} \\ \sum_{1} r_{012\ldots t} &\ldots k_{1t} \\ &\vdots \qquad\quad \vdots \\ \sum_{t-1} r_{012\ldots t} &= k_{(t-1)t} \end{aligned}$$

where k_{xt} = the number of individuals captured at time t who received a mark at time x. For example, it will be seen from the table given in the previous section that for the fourth sampling at $t = 3$,

$$\begin{aligned} k_{03} &= r_{03} + r_{013} + r_{023} + r_{0123}, \\ k_{13} &= r_{13} + r_{013} + r_{123} + r_{0123}, \\ k_{23} &= r_{23} + r_{023} + r_{123} + r_{0123}, \end{aligned}$$

the excess of the total number of marks over the number of marked individuals (s_3) being $r_{013} + r_{023} + r_{123} + 2r_{0123}$.

This system of grouping will be termed Method A.

Putting $\sum_{x=0}^{t-1} k_{xt} = s'_t$, the total number of previous marks found among the R_t individuals captured, and carrying out the same series of summations for the population numbers, we have the following modified table. (As before, the first column at each time t is the total number expected in the population as a whole, and the second is the number observed in the sample.)

Method A

$t =$	1		2		3		4	
	PR_0	k_{01}	P^2R_0	k_{02}	P^3R_0	k_{03}	P^4R_0	k_{04}
			PR_1	k_{12}	P^2R_1	k_{13}	P^3R_1	k_{14}
					PR_2	k_{23}	P^2R_2	k_{24}
							PR_3	k_{34}
Total	PR_0	s'_1	$P^2R_0+PR_1$	s'_2	$P^3R_0+P^2R_1+PR_2$	s'_3	$P^4R_0+P^3R_1+P^2R_2+PR_3$	s'_4

Evidently this method of grouping results in a considerable simplification of the expectation values. It may be shown, however, that in the long run it results in a loss of information.*

The shortest chain from which we may make an estimate of P consists of the first three samplings at $t = 0, 1$ and 2. From the above table we have for the relevant part of the log likelihood

$$L = k_{02}\log P - s'_2\log(PR_0+R_1),$$

and

$$-\frac{\partial^2 L}{\partial P^2} = \frac{k_{02}}{P^2} - \frac{s'_2 R_0^2}{(PR_0+R_1)^2}.$$

Replacing k_{02} by its expected value the total amount of information, yielded by the data for a three-point sampling grouped by Method A, is therefore

$$I_A = \frac{s'_2 R_0 R_1}{P(PR_0+R_1)^2}.$$

* Another way of considering this method of grouping is the following. At time t, R_t marks are returned to the population. Assuming a purely deterministic model, the number of these 't' marks which may be expected to survive until $t+a$ is P^aR_t. In this way the distribution of the expected number of marks at the time of some particular sample may easily be written down. We then assume that we have a random sample of this population of marks, and that the probability of obtaining the observed result is given by the appropriate term of a multinomial distribution. In order to estimate the unknown parameter P, we then take the combined likelihood of the series of results and form the maximum likelihood equations in the usual way.

It has been pointed out to us, however, that the actual sampling is not quite in accordance with this model. Each sample consists of individual animals, and any one animal may bear more than one mark. The marks in the different categories do not therefore appear to be free variables, since a number of individual marks may necessarily be associated together on the same animal at the time of sampling. Under these conditions is this use of the theory of maximum likelihood to obtain an asymptotic variance justifiable?

This is rather a complicated question on which we do not feel qualified to express an opinion. The difficulty arises when the number of recaptured animals bearing multiple marks becomes appreciable compared with those bearing only one mark. In practice, we should not, ourselves, adopt this method of grouping in such cases, but rather the system described later in the text as Method B. Nevertheless, we quite agree that there still remains this interesting theoretical question for consideration.

Taking the first three samplings as set out for the ungrouped data in the table given in the previous section, we have

$$L = r_{02}\log(PR_0 - r_{01}) - s_2\log(PR_0 + u_1),$$

and

$$-\frac{\partial^2 L}{\partial P^2} = \frac{r_{02}R_0^2}{(PR_0 - r_{01})^2} - \frac{s_2 R_0^2}{(PR_0 + u_1)^2}.$$

Thus, for the same set of data when ungrouped, the total amount of information is

$$I = \frac{s_2 R_0^2 R_1}{(PR_0 - r_{01})(PR_0 + u_1)^2}.$$

But $s_2' = s_2 + r_{012}$ and the expected value

$$E(r_{012}) = \frac{s_2 r_{01}}{PR_0 + u_1};$$

so that, for large samples,

$$E\left(\frac{s_2'}{s_2}\right) = \frac{PR_0 + R_1}{PR_0 + u_1}.$$

Hence

$$\frac{I_A}{I} = \frac{(PR_0 - r_{01})(PR_0 + R_1 - r_{01})}{PR_0(PR_0 + R_1)} < 1.$$

Thus, in the case of large samples, there will be a loss of information, even in the simplest chain, through grouping the results in terms of marks.

The loss of information in this simple case appears to arise owing to the fact that in grouping the observations in terms of marks, the expectations become functions merely of the total numbers caught at the previous samplings, whereas in terms of the different classes of marked individuals, they are functions both of the total numbers caught and of the frequencies observed in the various subclasses at previous samplings. Thus, in the latter case, the information that r_{01} individuals were recaptured at $t = 1$ is retained in the expectations for the sampling at $t = 2$, whereas it is lost in the former method of grouping. If this suggestion is correct, we might therefore expect that in a more extended chain of samplings the loss of information through grouping the observations in terms of marks would become progressively greater. In fact, it was partly in order to see whether this hypothetical increase in the loss of information was realized in practice that the sampling experiment described later was carried out.

4. A METHOD OF GROUPING INVOLVING NO LOSS OF INFORMATION

When the data are not grouped in any way, the expected numbers in the total population falling into the various subclasses of marked individuals are polynomial functions of the unknown P. Another method of grouping which immediately suggests itself is to sum the various classes at each sampling according to the degree of these polynomials. Thus it will be seen from the table given in §2 that at $t = 3$, for example, there is only one third-degree polynomial; but there is a pair of second degree which may be added together, and similarly for the four single-degree terms. Examining the suffixes of the observed numbers recaptured, it will be seen that this procedure is equivalent to grouping the recaptures according to the interval of time since they were last captured.

This system of grouping will be termed Method B.

If we define m_{xt} as the number of individuals captured at time t which were last captured at time x ($x = 0, 1, 2, \ldots, (t-1)$; $\sum_{x=0}^{t-1} m_{xt} = s_t$) we have the following table:

Method B

$t =$	1		2		3		4	
	PR_0	m_{01}	$P^2R_0 - Pm_{01}$	m_{02}	$P^3R_0 - P^2m_{01} - Pm_{02}$	m_{03}	$P^4R_0 - P^3m_{01} - P^2m_{02} - Pm_{03}$	m_{04}
			PR_1	m_{12}	$P^2R_1 - Pm_{12}$	m_{13}	$P^3R_1 - P^2m_{12} - Pm_{13}$	m_{14}
					PR_2	m_{23}	$P^2R_2 - Pm_{23}$	m_{24}
							PR_3	m_{34}
Total	PR_0	s_1	$P^2R_0 + Pu_1$	s_2	$P^3R_0 + P^2u_1 + Pu_2$	s_3	$P^4R_0 + P^3u_1 + P^2u_2 + Pu_3$	s_4

It may be shown that there is no loss of information through grouping the observations in this way. That is to say, defining the log likelihood for the ungrouped data as L, and that for the above table as L_B, if in $-\partial^2 L_B/\partial P^2$ and $-\partial^2 L/\partial P^2$ we substitute the expected value in each class for the corresponding observed value, then $I_B = I$.

It is easily seen that $-\partial^2 L/\partial P^2$ will differ from $-\partial^2 L_B/\partial P^2$ only in respect of those terms which are represented in the latter by the summation of the polynomials, the remaining terms being the same. Thus, for example, the only difference between the contributions to the total amount of information arising from the sampling at $t = 3$ will be that there is a term in

$$-\frac{\partial^2 L_B}{\partial P^2} = \ldots \frac{m_{13} R_1^2}{(PR_1 - m_{12})^2} \ldots,$$

corresponding to a pair of terms in

$$-\frac{\partial^2 L}{\partial P^2} = \ldots \frac{r_{13} u_1^2}{(Pu_1 - r_{12})^2} + \frac{r_{013} r_{01}^2}{(Pr_{01} - r_{012})^2} \ldots.$$

As a first step, if we replace the observed m_{12}, r_{12} and r_{012} by their corresponding expected values, we find that

$$\frac{R_1}{PR_1 - m_{12}} = \frac{u_1}{Pu_1 - r_{12}} = \frac{r_{01}}{Pr_{01} - r_{012}} = \frac{PR_0 + u_1}{P^2R_0 + Pu_1 - s_2};$$

and hence, since $E(r_{13}) + E(r_{013}) = E(m_{13})$, the contribution to the total amount of information will be the same in each case. By applying a similar type of argument to the remaining terms which differ in $-\partial^2 L/\partial P^2$ and $-\partial^2 L_B/\partial P^2$, it follows that $I_B = I$.

Thus, the method of grouping the recaptures according to the interval of time since they were last captured fulfils the requirements mentioned earlier. The subclasses which have to be considered are reduced in number, an obvious advantage in the case of a long chain of samplings, and in addition there is no loss of information. Thus, compared with the method of grouping by marks which is commonly employed, the present method should provide a more accurate estimate of P for a given set of data.

Various other methods of grouping the data were tried, but none of these proved so convenient as Method B. Either the maximum-likelihood equations were more troublesome to solve arithmetically, or there was a loss of information which varied in amount. One method is perhaps worth special mention, since, like Method A, it is attractive in so far as it

results in very simple expressions for the expected numbers. This is to group the recaptures according to the time they were first captured and marked. However, this, which we may call Method C, appears to result in a greater loss of information even than Method A. Thus, to take the simple three-point sampling as an example, we have

$$L_C = (r_{01}+r_{012})\log P - s_2 \log (PR_0+u_1),$$

and the expected number $$E(r_{01}+r_{012}) = \frac{s_2 PR_0}{PR_0+u_1},$$

from which we have in relation to the total amount of information for the ungrouped data,

$$\frac{I_C}{I} = \left(1-\frac{r_{01}}{R_1}\right)\frac{PR_0-r_{01}}{PR_0}.$$

Now, from the relationship given in § 3, we can write

$$\frac{I_A}{I} = \frac{I_A}{I_B} = \left(1-\frac{r_{01}}{PR_0+R_1}\right)\frac{PR_0-r_{01}}{PR_0},$$

and hence $$I_C < I_A < I_B.$$

5. Solution of the equations for estimating the survival factor

The solutions of the maximum-likelihood equations for the two methods of grouping the data are given here for a chain of five samplings. The way of extending the equations for a chain of greater length will be easily seen from the form of these solutions. In order to save space, the actual equations for L are not given, since these may be written down at sight from the tables in the previous sections.

A. *The method of grouping by marks*

$$\frac{\partial L}{\partial P} = \frac{1}{P}(k_{02}+2k_{03}+k_{13}+3k_{04}+2k_{14}+k_{24}) - \left[\frac{s_2' R_0}{PR_0+R_1}+\frac{s_3'(2PR_0+R_1)}{P^2R_0+PR_1+R_2}+\frac{s_4'(3P^2R_0+2PR_1+R_2)}{P^3R_0+P^2R_1+PR_2+R_3}\right] = 0.$$

In order to simplify the actual computations, this equation may be multiplied throughout by P and rewritten as

$$f(k) - \left[\frac{s_2' P^3 R_0}{P^3R_0+P^2R_1}+\frac{s_3'(2P^3R_0+P^2R_1)}{P^3R_0+P^2R_1+PR_2}+\frac{s_4'(3P^3R_0+2P^2R_1+PR_2)}{P^3R_0+P^2R_1+PR_2+R_3}\right] = 0,$$

where, setting out the observed distributions of marks in a convenient form of triangular diagram,

$$f(k) = k_{02}+2k_{03}+3k_{04} + k_{13}+2k_{14} + k_{24}.$$

It is not difficult to show by a little straightforward algebra that the form of solution illustrated by Fisher & Ford (1947) on several sets of observed data reduces to that given here. The estimate $\hat{P}$, together with its variance, may be obtained by iteration, and a numerical

example of the various steps in the solution is given in Appendix I. In solving this equation by the iterative method, however, it is most important to remember that for a trial value P_1, we are actually calculating the value of $P_1 \partial L/\partial P_1$. It is therefore advisable to obtain by division $\partial L/\partial P_1, \partial L/\partial P_2$, etc., before interpolating for the estimate $\hat{P}$ which makes $\partial L/\partial P = 0$. Although this correction may have only a trivial effect on the estimate $\hat{P}$, it becomes of importance in estimating the variance. For, if as a result of the iteration, we find two values, $P_1 < P_2$, between which the required value $\hat{P}$ lies, and if $\partial L/\partial P_1 = x$, say, and $\partial L/\partial P_2 = -y$, then we may take

$$\frac{1}{V(P)} = -\frac{\partial^2 L}{\partial P^2} \doteq \frac{y+x}{P_2 - P_1}.$$

By neglecting the fact that owing to the rearrangement of the original equation we are actually calculating $P\partial L/\partial P$ for trial values of P, we should therefore overestimate the variance unless this correction was made.

It will be seen that for the shortest possible chain, namely, a three-point sampling, the equation may be solved directly for P. Thus we have

$$\hat{P} = \frac{k_{02} R_1}{k_{12} R_0},$$

the variance of this estimate being (in terms of the observed values)

$$V(\hat{P}) = \frac{s_2' \hat{P}^2}{k_{02} k_{12}}.$$

But, for any chain of greater length, the iterative solution is the most convenient.

B. *The method of grouping the recaptures according to the interval of time since they were last captured*

Setting out the solution in the same triangular form as that used for the table of expected and observed values given in the previous section, we have

$$\begin{aligned}\frac{\partial L}{\partial P} = {} & \frac{m_{02} R_0}{PR_0 - m_{01}} + \frac{m_{03}(2PR_0 - m_{01})}{P^2 R_0 - Pm_{01} - m_{02}} + \frac{m_{04}(3P^2 R_0 - 2Pm_{01} - m_{02})}{P^3 R_0 - P^2 m_{01} - Pm_{02} - m_{03}} \\ & + \frac{m_{13} R_1}{PR_1 - m_{12}} + \frac{m_{14}(2PR_1 - m_{12})}{P^2 R_1 - Pm_{12} - m_{13}} \\ & + \frac{m_{24} R_2}{PR_2 - m_{23}} \\ & - \left[\frac{s_2 R_0}{PR_0 + u_1} + \frac{s_3(2PR_0 + u_1)}{P^2 R_0 + Pu_1 + u_2} + \frac{s_4(3P^2 R_0 + 2Pu_1 + u_2)}{P^3 R_0 + P^2 u_1 + Pu_2 + u_3} \right] \\ & = 0.\end{aligned}$$

Just as in the case of Method A, this equation, after a little rearrangement, may most easily be solved by iteration, and a numerical example of the solution is given in Appendix I. Although the type of calculation is very similar to that for Method A, the whole procedure takes longer. Compared with the time which must be spent in collecting the original data, however, any increase in time spent in computing is trivial, if a more accurate estimate of a particular parameter is obtained.

For the simple three-point sampling, we have the direct solution

$$\hat{P} = \frac{m_{01}s_2 + u_1 m_{02}}{m_{12}R_0},$$

and the variance (in terms of the observed values)

$$V(\hat{P}) = \frac{m_{02}s_2 R_1^2}{m_{12}^3 R_0^2}.$$

6. The design of the sampling experiment

It has been shown in §3 that the method of grouping defined here as Method A results, at any rate in the case of large samples, in a loss of information. This has been proved in the case of the shortest possible chain, namely, that given by the first three samplings at $t = 0$, 1 and 2, and it has been suggested that this loss of information might be expected to become greater as the length of chain increases. The question then naturally arises as to what we mean by 'large samples' in the case of a chain of such events. In the type of data which might be met with in practice, would there be in fact any advantage in adopting Method B instead of A?

It seems likely that in the type of field data such as that given by Fisher & Ford (1947) there would be very little to choose between the two methods, since their multiple-recapture rate was evidently very low and in this case the results would probably be much the same with either method. In such cases it may be preferable to group the data by Method A, simply owing to the greater ease with which the maximum-likelihood equation can be solved. But if the recapture rate be greater, there is an increased chance of multiple recaptures, and then the two ways of tabulating the data would become very different. Under these conditions the greater degree of precision expected in the estimates of the survival factor made from data grouped according to Method B might well be realized in practice.

Another closely allied problem of some practical interest is the following. Various species of small mammals, for example, the vole *Microtus agrestis*, pass the winter months in a non-breeding phase. It may be convenient, therefore, to carry out a series of live trappings during this period, since no dilution of the population is occurring through births. Suppose then that we wish to estimate the death-rate and numbers during such a period of, say, 5 months. Each sampling of the population requires not only a considerable expenditure of time and trouble over the purely technical side of live-trapping in the field, but also, in some cases, a special journey to the particular area studied. The question of the total cost of each sampling in relation to the amount of information obtained thus becomes important. Now, the minimum number of samples required in order to obtain an estimate of the death-rate is three. For a total period of 5 months we might therefore sample during every other month. Suppose that approximately the same number of animals can be captured and marked at each sampling, and that the cost in time, trouble and so forth, is constant, say x units per sample. If we increase the number of samples to five within the period, i.e. sampling at monthly intervals, the total cost increases in the ratio 5 : 3, or 1·67 times. Does the increase in the amount of information in regard to the various parameters which we wish to determine justify the additional labour and expense?

In order to investigate these points we decided to carry out an artificial sampling experiment in the laboratory. The scale of the experiment, the number and size of samples, and

the type of population were suggested by the analysis of some field data for the vole, *Microtus agrestis*, which were collected during the winter of 1937–8. On the area studied it was possible to live-trap and mark around fifty individuals at each sampling, and later calculations suggested that the size of this particular population at the beginning of the winter trappings was of the order of 300 individuals. The death-rate, on the average, appeared to be roughly 20 % per month. We were thus led to consider the following imaginary case.

A population of 300 individuals at some initial date decreases owing to deaths for a period of four units of time. A total of five samplings, including the initial one, are made at equidistant intervals, at each of which fifty individuals are withdrawn, marked and returned to the population. No births occur; immigration and emigration can be neglected; and a constant death-rate is assumed to be in operation during the whole period. The actual numbers alive (N_t) in the population at each of the successive samplings and the number of deaths (D_t) occurring during each interval, were taken to be:

t	N_t	D_t	P_t
0	300	48	0·8400
1	252	40	0·8413
2	212	34	0·8396
3	178	28	0·8427
4	150	—	—

The constant survival factor was assumed to be $P = 0{\cdot}84$, corresponding to an instantaneous death-rate per head of population $d = -\log_e P = 0{\cdot}194$, but owing to the necessity of having integral values of N_t, the actual values of P varied slightly for each interval. The average value over the interval of time 0–3, which is the most that can be determined from a total of five samplings, is $P = 0{\cdot}8403$. Since the estimates which are made later are expressed only to the third decimal place, the last figure being in error, we may take, for all ordinary purposes, the expected value as being $P = 0{\cdot}840$.

The individuals consisted of the counters which are made for the game of 'Lotto'. Each counter is a wooden, cylindrical disk of diameter 15 mm. and height 8 mm. with a number impressed on one of the two surfaces. They can be bought in numbered batches of ninety, and in order to distinguish different batches, they may be stained in dilute solutions of various coloured inks. Each individual counter can thus be defined by a colour and number. The shuffling and mixing of the counters before a sample was taken was done in a tin drum (diameter 15 cm., height 26 cm.) lined with felt and provided with an opening, closed by a sliding plate, on one side, for removing or returning the counters. This drum was fixed to a horizontal bar provided at one end with a turning handle and mounted in a wooden frame. This horizontal bar passed through the diagonal axis of the drum, so that when the drum was rotated the counters were also thrown from one end to the other. We are indebted to Mr D. A. Kempson for designing and making this apparatus.

Each chain of samplings was carried out in the following way. Starting with 300 counters which had been thoroughly mixed, a sample of fifty was taken, the individual counters being ticked off on a list according to their colour and number. This sample was returned and the total of 300 counters was again mixed by about twenty slow revolutions of the drum. Forty-eight counters were then taken at random and discarded, to represent the number of

deaths occurring during the first interval of time, and a further fifty counters were removed, ticked off on a list, and returned to the drum. The latter was again rotated to mix the counters, forty were removed to represent the deaths during the second interval, and a further fifty withdrawn for the sample at $t = 2$. This sequence of shufflings and withdrawals was continued until the fifth or last sampling of fifty counters at $t = 4$.

The whole chain of five samplings was repeated twenty times.

7. Preliminary analysis of the data

Before discussing the results of this experiment there are two points of some importance which must be mentioned. The first of these is the question whether the drawings were in fact random samples of the population. The chief difficulty we had in mind was whether there was a thorough mixing of such shaped counters after a reasonable number of revolutions—say about twenty—of the drum. Before embarking on the real experiment, therefore, a perfectly straightforward series of random samplings was made from a constant population. It is unnecessary to describe here in detail the analysis of these preliminary data; in all the tests which were applied the results were satisfactory, as judged by the usual statistical criteria. A further indication that the mixing and sampling were reasonably satisfactory is given by the results from the twenty replicate drawings from the population considered here.

Knowing the actual number of individuals in the population it is easy to calculate the expected number which should on the average occur in any particular class of marked individuals at each successive sample. These expected numbers are given in the first column of Table 1 for the data grouped according to Method B, and it will be seen from the second column that the observed means of the twenty replicates follow these very closely. (See also Appendix III, Table 3, where the results for the ungrouped data are given.) Moreover, if we include the unmarked class at each sample, there were four observed frequency distributions for each replicate chain, which could be compared with the expected frequencies by means of χ^2. When the data are grouped by Method B there are two classes at the first resampling, three at the second, and so on, so that the degrees of freedom are t for the appropriate distribution. The following were the sums of the twenty individual values of χ^2 at each sampling in the chain. None of these values is excessive; in fact, the agreement with expectation is, if anything, rather too good, but not exceptionally so, at the last sample. These results suggest very strongly that the shuffling of the counters between each drawing was satisfactory, since we might expect that any bias due to insufficient mixing would have a cumulative effect in such a chain.

t	χ^2	D.F. $= 20t$	$\sqrt{(2\chi^2)} - \sqrt{(40t-1)}$
1	19·43	20	—
2	48·35	40	+0·946
3	65·08	60	+0·499
4	67·39	80	−1·000

Although the values of χ^2 thus appear to be satisfactory, there is an important point which emerges when we consider the observed variance in each class of marked individuals. In the theoretical development of the first few sections we have adopted, in the first place, a purely deterministic model in which the death-rate falls equally on all subclasses of the population, and secondly, we have assumed that at each sampling the probability of getting the observed

result is given by a multinomial distribution. Thus, for example, at the first resampling when there are only two classes in the population at risk, the probability of observing m_{01} marked and u_1 unmarked individuals in a sample of R_1 is taken as

$$\frac{R_1!}{m_{01}!\,u_1!}\,p^{m_{01}}(1-p)^{u_1}.$$

Table 1. *The means and variances of the numbers falling into the various classes of marked individuals, when the data for the twenty replicates were grouped by Method B*

Class	Mean number of counters		Variance		
			Expected σ^2		s^2 (19 D.F.)
	Expected	Observed	Multinomial	Hypergeometric	
m_{01}	8·333	8·10	6·944	5·589	7·04
m_{02}	6·680	6·90	5·788	4·445	6·94
m_{12}	9·920	9·75	7·952	6·107	11·35
m_{03}	5·105	5·45	4·584	3·315	5·94
m_{13}	7·581	7·30	6·432	4·651	7·17
m_{23}	11·792	11·65	9·011	6·516	7·61
m_{04}	3·671	3·35	3·402	2·283	2·66
m_{14}	5·451	5·10	4·857	3·260	6·09
m_{24}	8·480	8·90	7·042	4·726	4·52
m_{34}	14·045	13·75	10·100	6·778	7·99

Now, strictly speaking, the appropriate distribution to be considered at each sampling is a hypergeometric of two or more classes. The error introduced by assuming a multinomial distribution is negligible when the size of the sample is small relative to the total number in the population sampled. But in this experiment the samples ranged from 17 to 33 % of the population at risk, and we should expect the observed variances to be closer to those expected for hypergeometric rather than multinomial distributions.

If p_i is the proportion of the population falling into the ith class, the expected number $E(m_i)$ in this class in a sample of R is for both types of distribution

$$E(m_i) = p_i R.$$

The expected variance of m_i is $\quad V(m_i) = p_i(1-p_i)\,R$

in the case of a multinomial distribution, while for the hypergeometric

$$V(m_i) = p_i(1-p_i)\,R\left(\frac{N-R}{N-1}\right),$$

where N is the total number in the population.

These expected variances are tabulated in the third and fourth columns of Table 1, the observed s^2 (19 D.F. in each case) being given in the last column. It will be seen that in nine cases out of ten the observed variance is greater than that expected for a hypergeometric distribution, and that in fact they approximate much more closely to the multinomial variances. It seems likely that this discrepancy arises owing to the way in which the experiment was carried out. Thus, to take as an example the first interval of time, the removal of

forty-eight counters chosen at random from the 300 in the drum to represent the number of deaths occurring during the interval introduces an additional random element. If there are $(N-R_0)$ unmarked and R_0 marked individuals, we should not expect to remove precisely $0{\cdot}16\,(N-R_0)$ and $0{\cdot}16\,R_0$ counters by this procedure. The assumption that at $t=1$ we are withdrawing a sample from a population in which the death-rate has fallen equally on the marked and unmarked class is thus not fulfilled. We really require some compound probability distribution for the expected numbers at the first resampling, and similarly for all the rest of the sequence. It might be expected that the general effect of this additional random element would be to increase the variance of the observed numbers occurring in the sub-classes at each sampling, and it is suggested here that the excess of the observed variances in Table 1 over those expected for a hypergeometric distribution is due to this cause.

The point raised here may be of considerable practical importance, since in the case of a biological population the death-rate during some interval of time must be regarded as a probability rather than as a constant parameter. Strictly speaking, some appropriate form of stochastic model is needed to describe the changes in the population and the expected distribution of marked individuals at each sampling. The chief difficulty which may arise from the use of a deterministic model is in regard to the variances of any estimates which we may make. If an estimate of the variance of some parameter is derived from the maximum-likelihood equation for this type of model, it is possible that this may be an underestimate of the true sampling variance, more especially when the number of recaptures is small. Although, as will be seen from the results given in the next section, the variances observed between our twenty replicates are very close to those expected theoretically under the assumptions which have been made, this agreement may only be due to the scale on which this experiment was carried out. For smaller samples, or for a smaller proportion of the population sampled, the results might not have been so satisfactory. However, we shall be returning to this point in a later section.

8. Results of the sampling experiment

The following figures for the experimental population show how the total amount of information in regard to P should increase with the length of the sampling chain. These values of I_P were obtained by inserting the expected numbers in the appropriate equations for $\partial^2 L/\partial P^2$. Thus, the amount of information increases very rapidly with the length of chain, and the

Chain ending at t	I_P for data grouped by		100 A/B
	Method A	Method B	
2	6·41	8·79	73
3	34·34	53·74	64
4	110·38	203·29	54

loss of efficiency through grouping the data by Method A becomes progressively greater. The efficiency of A is only 54 % for a chain of five samplings. Naturally, it is difficult to make any general statements from figures which are based on a specific numerical example, but roughly we should expect much the same comparative degree of efficiency in any non-breeding population of this type, provided that our samples remained around 20 % of the individuals at risk.

The further problem considered in §6 was: given a limited period of time T between the first and last samples in a chain, whether the increase in the amount of information through increasing the number of samples from the minimum of three, to a total of five, is sufficient to justify the additional labour and expense. In the present example we obtained the following expected figures.

No. of samples for the period 0–T	I_P		$\pm 100\sigma/P$	
	Method A	Method B	A	B
3	27·64	40·09	± 22·6	± 18·8
5	110·38	203·29	± 11·3	± 8·4

Thus, taking B as the most efficient method of grouping, the average amount of information per sample should increase from $40/3 = 13{\cdot}3$ to $203/5 = 40{\cdot}6$, or roughly three times. If the average cost of each sampling remains constant, the two additional samples are therefore most repaying. We should expect the percentage error of our estimate of the survival factor ($P = 0{\cdot}84$) to decrease from about 19 to 8·4, which is quite a marked increase in precision.

It is interesting to see how far these expected figures were realized in the sampling experiment. The estimates of P are given in Table 2 for each of the twenty replicate drawings A, B, C, ..., S, T.

Three-point sampling

In this series only the samples taken at $t = 0$, 2 and 4 were considered, and in the first two columns of Table 2 the estimates $\hat{P}$ are expressed per unit of time. (The values of $\hat{P}^2$ over the period 0–2 were calculated directly from the equations for a three-point sampling given in §5, pp. 277 and 278.) It will be seen that the two methods of grouping gave very similar results. On twelve occasions out of twenty the estimate from Method B was nearer the true value of $P = 0{\cdot}84$; but the means and variances for the twenty replicates were almost identical. The two means are slightly below the true value, but the difference is not particularly marked, since the standard error is in both cases of the order $\pm 0{\cdot}035$. Evidently there is little to choose between the two methods of grouping, as judged by these data, in the case of a three-point sampling. There is no evidence of any marked bias in the estimates, and the differences between the observed and expected variances are no greater than those which might easily occur in random samples of twenty. But, although in one sense these agreements with expectation are satisfactory, the general run of the results is not very promising from the point of view of the biologist. Actually, in three drawings out of twenty, we obtained an estimate of P which was greater than unity. If any one of these three drawings happened to be the result of a single experiment in the field, this would lead to the nonsensical conclusion that a negative death-rate was in operation.* In other cases we should grossly overestimate the death-rate. Clearly, a three-point sampling of around 20 % of a population consisting

* It is to be noted that in the mathematical development no definition of the symbol P has been made restricting it to values $\leqslant 1$. Exactly the same equations would arise in the imaginary case of a population increasing, for instance, by division, in which any parent possessed the peculiar property of handing on to the offspring the same combination of marks it happened to have at the moment of division.

of a few hundred individuals is not a very efficient method of estimating this parameter. Nevertheless, as we hope to show in a later paper, the results of such a chain may still be used to obtain an estimate of the total numbers, which will very often be sufficiently accurate, if we are interested merely in the order of magnitude of the population sampled.

Five-point sampling

The estimates of P were obtained in both series by the iterative solution of the equations given in §5. In the case of Method A, the solutions in this experiment were greatly simplified by the fact that the number of individuals marked was the same at each sampling. Hence, as will be seen by putting $R_0 = R_1 = \ldots = R_T$ in the appropriate equation, a series of tables for trial values of P could be prepared. (This property of the equation for estimating P from the data grouped by Method A may possibly be of use in the design of field experiments, whenever the size of each sample can be fixed in advance and a large amount of data is to be collected and analysed.) No great accuracy in the solution was attempted, and an interval of 0·05 in the argument was adopted for the tables, the final estimate being made by simple linear interpolation between the two nearest trial values. No tables could be prepared for the solution of the equation for Method B. Generally speaking, the same interval of 0·05 between trial values of P was adopted in the iteration, though occasionally if the first value tried gave an answer very close to the solution, a smaller interval would be chosen. The estimates of $I_P = -\partial^2 L/\partial P^2$ obtained from the iteration are given in Table 2 for both series, the reciprocals of these values being, of course, the estimated variance in each case. Although no great accuracy was attempted in the iteration, it is interesting to see that, in both series, the mean of these estimates of I_P is very close to the expected value.

It will be seen from the results given in Table 2 that the expected difference between the two methods of grouping is more fully realized in the five-point sampling. In only three cases out of twenty was the estimate of P from the data grouped by Method A closer to the true value of $P = 0{\cdot}84$, and in one of these cases the difference between the two methods was trifling. On the average the estimates for Method B were nearer the true value by an amount equal to $0{\cdot}0171 \pm 0{\cdot}0072$. This greater degree of precision is represented in the observed variances between replicates, which were $s^2 = 0{\cdot}004227$ for Method B, compared with $s^2 = 0{\cdot}007647$ for Method A. Thus, judging from these results, the efficiency of A would be reckoned as about 55 %, a figure which is remarkably close to the theoretical one of 54 %.

Both of the observed variances are somewhat less than those expected. The discrepancy, however, is not at all marked and is of a degree which might very well arise in a random sample of twenty observations ($\chi^2 = 16{\cdot}0$ and $16{\cdot}3$ for series A and B respectively, there being 19 D.F.). Out of interest, we also calculated the expected variance, when allowance was made for the size of the samples relative to the population at risk, by multiplying the contribution to $\partial^2 L/\partial P^2$ at each sampling by the appropriate factor $(N_t - 1)/(N_t - R_t)$. It was found in this way that for the data grouped according to Method B the expected $\sigma^2 = 0{\cdot}003373$, a figure not very much below that of $\sigma^2 = 0{\cdot}004919$ for an infinite population. The observed variance between replicates, $s^2 = 0{\cdot}004227$, is compatible with either value.

The results of this experiment, therefore, bear out the theoretical expectations. In sampling for a limited period of time on this scale from a population of a few hundred individuals, the increased precision obtained through increasing the length of the sampling chain is apparent. It is evident, too, that the method of grouping the recaptures according to the interval of time since they were last captured is preferable to that of grouping by marks. Under

conditions similar to those in the experiment, the error of an estimate $\hat{P}$ is thereby reduced to around 8 % for a chain of five samples. The agreement between the observed inter-replicate variances, both for the three- and the five-point samplings, with those expected for large samples from a deterministic model population is of particular interest. Keeping constant the proportion of the population sampled, and allowing the size of the samples (R) to increase, we should therefore expect the variance of an estimate $\hat{P}$ to be $50/R$ times the values given here. Thus, the variance would be halved for samples of 100 from a similar population of twice the size. The behaviour of the variance, if we were either to decrease the size of the samples, or to sample a smaller proportion of the population at risk, needs further consideration however. We may inquire what the results of this experiment would have been if we had taken a much smaller number of counters at each sampling of the population. Would the variance between replicates then have been in so close agreement with expectation?

Table 2. *Estimates of the survival factor P from the data grouped by Methods A and B*

(Expected value, $P = 0{\cdot}840$)

Replicate	Three-point sampling		Five-point sampling			
	Method A	Method B	Method A		Method B	
	$\hat{P}$	$\hat{P}$	$\hat{P}$	$I_{\hat{P}}$	$\hat{P}$	$I_{\hat{P}}$
A	0·775	0·821	0·744	161	0·774	263
B	0·620	0·609	0·725	138	0·784	263
C	0·798	0·702	0·809	93	0·874	199
D	0·782	0·665	0·995	105	0·912	211
E	0·646	0·719	0·774	116	0·816	199
F	0·832	0·858	0·931	89	0·917	127
G	1·080	1·161	0·816	101	0·849	284
H	0·853	0·784	0·852	98	0·847	196
I	0·866	0·898	0·899	100	0·872	153
J	0·935	0·930	0·845	114	0·884	223
K	0·707	0·733	0·697	131	0·699	230
L	1·069	1·058	0·863	91	0·862	175
M	0·707	0·721	0·811	98	0·837	201
N	0·798	0·710	0·834	116	0·836	192
O	0·598	0·602	0·793	138	0·798	241
P	0·598	0·674	0·721	137	0·738	236
Q	0·961	0·869	0·994	86	0·942	189
R	0·775	0·792	0·907	93	0·884	179
S	0·817	0·850	0·708	144	0·751	303
T	1·118	1·107	0·843	103	0·904	176
Mean	0·8131	0·8167	0·8280	112·6	0·8390	212·0
s^2 (19 D.F.)	0·02397	0·02475	0·007647	—	0·004227	—
Expected σ^2 or I_P	0·03618	0·02494	0·009058	110·4	0·004919	203·3

9. THE EFFECT OF REDUCING THE SIZE OF SAMPLE

An attempt was made to answer this question from the data already collected in the experiment, by supposing each sample to have been half the original size. As we have indicated

in §6, the individual counters forming each sample of fifty were ticked off on a list of the whole 300 counters used in the experiment. These lists were gone through, and the entry for every alternate counter in each sample was marked with a circle. Considering in the first place only the entries marked with a circle, we then had a chain of five samples of twenty-five for each of the original twenty replicates, and the results could be tabulated in any way which seemed the most convenient. Similarly, by considering the entries which were not marked with a circle, we could obtain another set of twenty chains. For each of the replicates A, B, C, ..., T, we therefore had a pair of chains, each consisting of five samples of twenty-five, from the original population.

The recaptures in each chain were then grouped by Method A, simply in order to reduce the labour of estimating the forty values of P. For a five-point sampling, in which each sample consisted of twenty-five counters, the expected number of marks in the various k_{xt} classes for this method of grouping ranged from 2·08 to 3·51, compared with values four times as great in the case of the original samples of fifty from this population. Thus, by halving the size of sample, we have reduced the expectations to relatively small numbers. The expected variance of P from the maximum-likelihood equation was $\sigma^2 = 0{\cdot}0362$, which is the same as the expected variance for a three-point sampling of fifty for the original data grouped by Method A (see §8, Table 2).

The mean of the forty estimates was $\overline{P} = 0{\cdot}864$. We then assumed that we had a pair of independent estimates for each of the twenty replicates (the correlation between the twenty pairs was $r = +0{\cdot}10$) and carried out an analysis of variance in the usual way. The following were the results, together with the values of χ^2 obtained by contrasting each mean square with σ^2:

	D.F.	Mean square	χ^2	P
Between replicates	19	0·06632	34·8	$<0{\cdot}02>0{\cdot}01$
Within replicates	20	0·05412	29·9	$<0{\cdot}10>0{\cdot}05$
Total	39	0·06006		

These variances are greater than $\sigma^2 = 0{\cdot}0362$; indeed, the discrepancy for the 'between replicates' s^2, which is the comparable variance to those calculated in the previous section, would be reckoned as significant. In addition, the values of $I_{\hat{P}}$ were calculated from the iteration, and the reciprocal of the mean of these forty values was 0·0337. On the average, therefore, the individual estimates of the variance tended to be less than the observed variance between replicates.

The results of this experiment to investigate the effect of reducing the size of sample are in no way decisive. They suggest, if nothing more, that further work is needed on this problem, which is one both of theoretical and practical importance. It would be extremely interesting to know whether these discrepancies between observation and expectation could be attributed to the use of a deterministic, instead of a stochastic, model to represent the sampling chain, or to the fact that the expected numbers in the various classes were small; or, finally, whether they should be regarded merely as a chance occurrence of the somewhat rare type of event which must always be reckoned with as a possibility in a sampling experiment of this nature.

10. SUMMARY

The problem considered in this paper is that of estimating the death-rate in a population which is continuously changing in numbers, from the results of a chain of successive samples at each of which a number of individuals are captured, marked and returned to the population (the capture-recapture method). In order to solve this problem in a comparatively straightforward fashion, certain simplifying assumptions must be made, of which all may not be realized in practice. In the particular case studied, the death-rate Q per unit of time, or alternatively the survival factor $P = 1-Q$, in the population as a whole, is assumed to remain constant over the period of sampling.

At each resampling of the population the history of the marked individuals is found from the marks received at previous captures. Since the number of possible classes of previous markings increases very rapidly with the length of the sampling chain, it is convenient to adopt some method of grouping the observations, and the efficiency of various methods is discussed in relation to the total amount of information yielded by the data. It is shown that the method of grouping which consists in forming the distributions of marks observed among the recaptured animals, and which is termed here Method A, results in a loss of information. Nevertheless, this method may still be the most convenient when the proportion of recaptures in the samples is low and a large amount of data has to be analysed. Another method, termed here Method B, in which the recaptures are grouped according to the interval of time since they were last captured, was found to have certain advantages, of which the most important is that there is no loss of information. Compared with Method A, therefore, a more accurate estimate of the survival factor should be obtained from a given set of data grouped in this way. The maximum-likelihood equations are given for estimating the survival factor from the data grouped by these two methods, and numerical examples of the solutions are also given in an appendix.

In order to see whether the comparative efficiency of these two methods of grouping would be realized in the type of data which might be met with in practice, a sampling experiment was carried out in the laboratory. The scale of the experiment and the type of population were suggested by the analysis of some field data. The artificial population, which initially consisted of 300 numbered counters, was assumed to be subject to a constant 'death-rate', and five samples, each of fifty counters, were taken at equidistant intervals of time. The whole chain of five samplings was repeated twenty times.

The results of the main experiment were in close agreement with those expected theoretically from the maximum-likelihood equations for a deterministic model of the population. They are also used to show how the total amount of information in regard to the death-rate increases very rapidly according to the number of samples taken within a limited period of time, a point which is of importance in relation to the cost and design of investigations in the field. Using the data collected in this experiment, an attempt was made to investigate the effect on the variance between replicates of reducing the expected numbers in the various classes of recaptures to relatively small numbers. The results were inconclusive, but were sufficient to indicate that further work is needed on this question, which is one not only of theoretical, but also of practical interest.

The expenses of this investigation were partly paid for by a grant from the Agricultural Research Council.

APPENDIX I

Numerical examples of the solutions of the equations

The iterative solutions of the maximum-likelihood equations for the two methods of grouping the data are illustrated here on the same numerical example (replicate M).

Method A

Writing the observed k_{xt} in a form of triangular diagram, which is a convenient way of exhibiting the data, the equation to be solved is

$$\begin{aligned} & k_{02}+2k_{03}+3k_{04} \\ & \quad + k_{13}+2k_{14} \\ & \quad\quad + k_{24} \\ -\Big[\frac{s'_2P^2R_0}{P^2R_0+P^2R_1} & + \frac{s'_3(2P^2R_0+P^2R_1)}{P^2R_0+P^2R_1+PR_2} + \frac{s'_4(3P^2R_0+2P^2R_1+PR_2)}{P^2R_0+P^2R_1+PR_2+R_3}\Big] \\ = 0. & \end{aligned}$$

It will be seen that the positive term in this equation is merely a function of the observations and can be calculated immediately. In order to obtain the quotients for multiplying the observed sums of marks (s'_t), we may proceed as follows. (The same type of calculation also occurs in the solution of the equation for Method B.) Choose some trial value of P and write down the required powers of P in a column, starting with the highest power. Opposite these enter the appropriate values of R_t and form the products. The remaining steps may be set out as follows:

	S_1	S_2
P^3R_0	P^3R_0	
P^2R_1	$P^3R_0+P^2R_1$	P^3R_0
PR_2	$P^3R_0+P^2R_1+PR_2$	$2P^3R_0+P^2R_1$
R_3	$P^3R_0+P^2R_1+PR_2+R_3$	$3P^3R_0+2P^2R_1+PR_2$

The entries in the first column are summed from above down to give the second or S_1 column. This is in turn summed, again from the top, to give S_2, the first term being entered one line lower down and so on. Each entry in the S_2 column is then divided by the appropriate entry in S_1, giving the quotients by which the observed s'_t must be multiplied. For machine work it is not necessary to write down the products of the powers of P with the number of captures, provided the same number of decimal places is retained, since the first summation can be carried out directly on the machine.

The following are the recorded results (k_{xt}) for replicate M, together with the various totals which are required in the calculations:

x \ t	0	1	2	3	4
0	.	9	6	10	5
1	.	.	7	7	10
2	.	.	.	15	10
3	.	.	.	.	11
s'_t	.	.	13	32	36
R_t	50	50	50	50	.

Then $$f(k) = 6+2\times 10+3\times 5+7+2\times 10+10 = 78.$$

Taking the trial value, $P_1 = 0{\cdot}8$, we have

	S_1	S_2	S_2/S_1	s'_t
$P_1^3 = 0{\cdot}512$	25·600			
$P_1^2 = 0{\cdot}640$	57·600	25·600	0·4444	13
$P_1 = 0{\cdot}800$	97·600	83·200	0·8525	32
1	147·600	180·800	1·2249	36

Then $$13 \times 0{\cdot}4444 + 32 \times 0{\cdot}8525 + 36 \times 1{\cdot}2249 = 77{\cdot}154,$$

and subtracting this quantity from 78, the value of $P_1 \partial L/\partial P_1$ for $P_1 = 0{\cdot}8$ is $+0{\cdot}846$.

Since this quantity is positive, the required value of P is greater than 0·8. Taking a new trial value $P_2 = 0{\cdot}85$, and repeating the summations, we have

$$P_2 \partial L/\partial P_2 = 78 - 81{\cdot}263 = -3{\cdot}263.$$

Hence, for $P_1 = 0{\cdot}8$ and $P_2 = 0{\cdot}85$,

$$\frac{\partial L}{\partial P_1} = +1{\cdot}058, \quad \frac{\partial L}{\partial P_2} = -3{\cdot}839.$$

Thus, by interpolation, we have as a first approximation,

$$\hat{P} = 0{\cdot}8 + \frac{1{\cdot}058 \times 0{\cdot}05}{4{\cdot}897} = 0{\cdot}811;$$

and

$$I_P = 1/V(P) \doteqdot 4{\cdot}897/0{\cdot}05 = 98;$$

whence the estimated variance of $\hat{P}$ is 0·01020, and the standard error $s = \pm 0{\cdot}101$.

If a more accurate solution is needed, the value $P = 0{\cdot}811$ is then substituted, and so on. The method of obtaining a more accurate estimate of I_P, given the results for three trial values of P, is illustrated by Fisher (1950, §57·2).

Method B

In order to obtain a form more suitable for the calculations, the equation for $\partial L/\partial P = 0$ in §5 is first multiplied throughout by P and rewritten as

$$\begin{aligned}
&\frac{m_{02}P^3R_0}{P^3R_0 - P^2m_{01}} + \frac{m_{03}(2R^3R_0 - P^2m_{01})}{P^3R_0 - P^2m_{01} - Pm_{02}} + \frac{m_{04}(3P^3R_0 - 2P^2m_{01} - Pm_{02})}{P^3R_0 - P^2m_{01} - Pm_{02} - m_{03}} \\
&\qquad + \frac{m_{13}P^2R_1}{P^2R_1 - Pm_{12}} \qquad + \frac{m_{14}(2P^2R_1 - Pm_{12})}{P^2R_1 - Pm_{12} - m_{13}} \\
&\qquad\qquad\qquad\qquad + \frac{m_{24}PR_2}{PR_2 - m_{23}} \\
&= \frac{s_2P^3R_0}{P^3R_0 + P^2u_1} + \frac{s_3(2P^3R_0 + P^2u_1)}{P^3R_0 + P^2u_1 + Pu_2} \quad + \frac{s_4(3P^3R_0 + 2P^2u_1 + Pu_2)}{P^3R_0 + P^2u_1 + Pu_2 + u_3}.
\end{aligned}$$

Looking at the three rows of the triangular form in which the left-hand side (L.H.S.) of the equation is written, it will be seen that the six quotients involving the powers of P may be obtained by three series of double summations in the same way as the quotients in the previous example. Similarly, the three quotients in the right-hand side (R.H.S.). The actual calculations are thus of the same form, apart from the question of signs, as in the case of Method A, but the whole process is lengthier.

Taking the same replicate M as an example, we have the following table of the numbers observed in the different classes (m_{xt}), together with the number of unmarked individuals (u_t) which will be required in the calculations:

x \ t	0	1	2	3	4
0	.	9	6	7	2
1	.	.	7	5	6
2	.	.	.	15	7
3	.	.	.	.	11
s_t	.	.	13	27	26
u_t	50	41	37	23	.
R_t	50	50	50	.	.

Taking $P_1 = 0{\cdot}8$ as the first trial value, we have the following summations for the three rows of the L.H.S. It will be seen from the equation that in each case the first summation column involves a series of subtractions, and the second summation a series of additions.

	S_1	S_2	S_2/S_1	m_{xt}
$P_1^3 = 0{\cdot}512$	25·600			
$P_1^2 = 0{\cdot}640$	19·840	25·600	1·2903	6
$P_1 = 0{\cdot}800$	15·040	45·440	3·0213	7
1	8·040	60·480	7·5224	2
$P_1^2 = 0{\cdot}640$	32·000			
$P_1 = 0{\cdot}800$	26·400	32·000	1·2121	5
1	21·400	58·400	2·7290	6
$P_1 = 0{\cdot}800$	40·000			
1	25·000	40·000	1·6000	7

Hence, by adding the cross products of the last two columns, the L.H.S. = 77·570.

The calculations for the R.H.S. are similar, except that only additions are involved.

	S_1	S_2	S_2/S_1	s_t
$P_1^3 = 0{\cdot}512$	25·600			
$P_1^2 = 0{\cdot}640$	51·840	25·600	0·4938	13
$P_1 = 0{\cdot}800$	81·440	77·440	0·9509	27
1	104·440	158·880	1·5213	26

and the R.H.S. = 71·647.

We have, therefore, for $P_1 = 0{\cdot}8$, $P_1 \partial L/\partial P_1 = +5{\cdot}923$. (It is important to remember that in rearranging the original equation for $\partial L/\partial P = 0$, we multiplied throughout by P.) Since this quantity is positive, a larger value of P, say $P_2 = 0{\cdot}85$, is tried. Carrying out the same series of summations, etc., we find $P_2 \partial L/\partial P_2 = -2{\cdot}240$. Thus, for $P_1 = 0{\cdot}8$, and $P_2 = 0{\cdot}85$,

$$\frac{\partial L}{\partial P_1} = +7{\cdot}404, \quad \frac{\partial L}{\partial P_2} = -2{\cdot}635,$$

and for our first approximation,

$$\hat{P} = 0{\cdot}80 + \frac{7{\cdot}404 \times 0{\cdot}05}{10{\cdot}039} = 0{\cdot}837.$$

Also $$I_P \doteqdot 10{\cdot}039/0{\cdot}05 = 201, \quad V(P) = 0{\cdot}004975 \quad \text{and} \quad s = \pm 0{\cdot}071$$

APPENDIX II

Corrections to allow for accidental deaths

It has been assumed hitherto that the number of individuals captured at each sampling (C_t) is equal to the number marked and returned to the population (R_t). In practice, however, it occasionally happens that a certain number of accidental deaths occur owing to the method of capture employed. Thus, some animals may be found dead in the traps, or through some accident in handling are so damaged that it is necessary to destroy them. Again, a group of individuals may be removed from the population for some purpose or other. It may be necessary, therefore, to allow for such losses at each sampling in estimating the survival factor.

In the case of Method B, the corrections which must be made in the equation given in Appendix I are extremely simple. The number of unmarked animals (u_t) and the number of recaptures falling into the various classes (m_{xt}: $\sum_x m_{xt} = s_t$) will be based, naturally, on the total number of captures, dead and alive, at time t. Let D_t be the number of accidental deaths and removals at time t, so that the number marked and returned alive to the population is $R_t = C_t - D_t$. Then, if we write out the expected numbers in the population as a whole at each sampling, in a similar form to that given in § 4, it will be found that the maximum-likelihood equation for estimating P is the same as that given in Appendix I, except that the $u_t (t = 1, 2, 3)$ in the R.H.S. are replaced by quantities which we shall define as

$$y_t = u_t - D_t;$$

namely, the number of unmarked individuals captured, both dead and alive, less the total number of accidental deaths and removals. Thus, to summarize the corrections: in the equation for Method B, base

Table 3. The number of recaptures observed in the various classes

Replicate

Table 3. *The number of recaptures observed in the various classes*

Replicate

Suffix of the recapture class at time t ($r_{012\ldots t}$)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	Mean	
																					Observed	Expected
01	9	7	12	13	13	7	8	9	9	8	8	6	9	7	6	2	8	6	5	10	8·10	8·333
02	5	5	11	10	5	5	8	9	5	11	4	5	6	5	10	12	7	6	5	4	6·90	6·680
12	15	10	6	3	6	7	13	5	8	11	10	8	7	8	10	7	7	9	13	9	8·60	8·267
012	2	2	—	1	—	2	1	1	2	1	2	1	—	2	1	—	—	2	2	1	1·15	1·653
s_2	22	17	17	14	11	14	22	15	15	23	16	14	13	15	21	19	14	17	20	14	16·65	16·600
03	3	6	3	10	4	4	8	5	5	4	1	7	7	5	8	1	8	7	5	8	5·45	5·105
13	2	8	6	4	9	7	3	3	4	8	3	8	3	8	7	7	7	9	7	6	5·95	6·317
23	9	10	8	7	7	5	4	13	8	9	8	9	12	11	6	8	6	7	7	7	8·05	7·877
013	2	1	2	5	4	1	1	1	—	2	—	1	2	1	1	—	1	2	—	—	1·35	1·263
023	—	—	1	5	2	1	2	1	3	2	1	—	1	—	1	2	1	1	2	1	1·35	1·576
123	3	—	2	1	1	—	—	1	4	2	2	2	2	3	3	1	3	1	4	2	1·85	1·950
0123	1	—	—	—	—	—	1	—	1	1	2	—	—	1	—	—	—	—	1	—	0·40	0·390
s_3	20	25	22	32	27	18	19	24	25	28	17	27	27	29	26	19	26	27	26	24	24·40	24·478
04	7	2	3	1	3	4	3	2	4	4	5	4	2	2	—	3	4	4	4	6	3·35	3·671
14	2	4	4	2	2	8	3	7	5	4	2	4	6	8	6	6	3	4	3	3	4·30	4·543
24	7	5	2	6	7	8	3	5	9	2	7	4	6	5	6	6	7	7	7	6	5·75	5·665
34	10	9	8	4	9	7	12	5	4	8	9	8	4	5	11	9	7	4	12	9	7·70	7·168
014	—	—	—	—	1	2	—	2	3	1	—	1	—	1	1	—	3	—	1	—	0·80	0·908
024	—	1	4	2	—	1	—	2	—	1	1	1	—	3	2	2	4	2	1	1	1·40	1·133
034	—	1	—	1	1	2	3	1	1	—	—	2	2	1	1	—	1	3	2	3	1·25	1·434
124	4	2	1	2	1	2	2	1	1	2	3	2	1	1	2	1	1	3	1	—	1·65	1·402
134	1	1	2	1	1	1	2	1	1	2	1	3	2	3	1	3	2	2	1	2	1·65	1·774
234	2	4	4	3	3	2	1	3	1	2	1	—	1	2	2	4	1	3	3	—	2·10	2·213
0124	—	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·10	0·281
0134	1	—	—	2	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	0·20	0·355
0234	—	—	—	4	—	—	—	1	1	—	—	—	1	—	1	—	—	—	—	—	0·40	0·443
1234	1	—	—	—	1	—	—	1	—	—	—	—	1	—	1	1	—	—	—	1	0·35	0·547
01234	1	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	0·10	0·110
s_4	36	30	28	29	29	37	30	31	30	27	29	29	26	31	34	35	33	32	35	31	31·10	31·647

the m_{st} and the s_t on the numbers captured, the R_t on the numbers returned alive to the population, and finally replace all u_t by y_t.

In the case of Method A, the full corrections which should be made are more complicated. Thus, it is necessary to form each time the distribution of marks observed among the accidental deaths and removals. These values must be subtracted from the expected population numbers in the appropriate classes at each sampling. The full equation which results is very similar in its general form to that for Method B, and thus the comparative simplicity of the original solution is lost. Owing to the loss of information which results from grouping the data by this method, it therefore seems simpler to adopt Method B in the first place and apply the corrections given above. Nevertheless, if the number of accidental deaths is not large, or the recaptures form only a small proportion of each catch, we may still group the recaptures by Method A and retain the numbers released alive (R_t) in the equation. This is equivalent to assuming that the accidental deaths occur among the unmarked animals at each sampling. This procedure will probably be accurate enough for all ordinary purposes, provided the number of losses is small.

APPENDIX III

Detailed results of the sampling experiment

The detailed results of the sampling experiment are given in Table 3, where the observed recaptures $r_{012\ldots t}$ at time t are tabulated, for each of the twenty replicate chains A, B, C, ..., S, T, according to the $n = 2^t - 1$ possible classes of suffix $012\ldots t$. The total number of counters in the population at $t = 0, 1, 2, 3, 4$ was 300, 252, 212, 178 and 150 respectively, and each sample consisted of fifty counters. The method of carrying out each chain of drawings is described in § 6. Only the total number of previously marked counters (s_t) recovered at time t is given in the table, since the number unmarked can easily be found by subtraction, $u_t = 50 - s_t$. The last two columns give the observed means and those expected from the deterministic model of the population.

REFERENCES

DOWDESWELL, W. H., FISHER, R. A. & FORD, E. B. (1940). The quantitative study of populations in the Lepidoptera. I. *Polyommatus icarus* Rott. *Ann. Eugen., Lond.*, **10**, 123–136.

FISHER, R. A. (1950). *Statistical Methods for Research Workers*, 11th ed. Edinburgh: Oliver and Boyd.

FISHER, R. A. & FORD, E. B. (1947). The spread of a gene in natural conditions in a colony of the moth, *Panaxia dominula* L. *Heredity*, **1**, 143–74.

JACKSON, C. H. N. (1939). The analysis of an animal population. *J. Anim. Ecol.* **8**, 238–46.

JACKSON, C. H. N. (1948). The analysis of a tsetse-fly population. III. *Ann. Eugen., Lond.*, **14**, 91–108.

ON ESTIMATING THE SIZE OF MOBILE POPULATIONS FROM RECAPTURE DATA

By NORMAN T. J. BAILEY, *Department of Human Ecology, University of Cambridge*

CONTENTS

		PAGE
1.	Introduction	293
2.	Simple recapture	294
3.	Jackson's 'negative' and 'positive' methods	298
4.	Trellis diagrams	301
5.	The 'triple-catch' method	302
6.	Summary	305
	References	306

1. Introduction

The estimation of the total size of plant or animal populations is of great importance in a variety of biological problems, which may relate to population growth, ecological adaptation, genetic constitution, natural selection and evolution, and so on. Obvious practical consequences are the maintenance of human food supplies and the control of insect pests. For human communities and populations of sessile organisms procedures employing fixed sampling units are available, but for mobile populations other methods must be used.

The basic technique, which seems to have been first used by Lincoln (1930) to estimate the total number of duck in North America, and is sometimes referred to as the 'Lincoln Index', is as follows. One catches, marks and releases a certain number of animals taken at random from the population. A further random sample is caught and the proportion of marked animals noted. Then the *total* number of marked animals released divided by the proportion of marked animals in the sample captured can clearly be used as an estimate of the total population size. The same method was adopted independently by Jackson (1933), who used it for estimating the true density of tsetse flies. Jackson (1937, 1939) subsequently extended his treatment to allow for both birth- and death-rates, advocating his 'negative' and 'positive' methods. Further work by Jackson (1940) took into account the effects of migration. Recapture methods have also been successfully applied to populations of Lepidoptera. Dowdeswell, Fisher & Ford (1940) caught, marked and released moths on several different days. On each day the marked insects were classified according to the day on which they had previously been released. Analysis of the data, which could be exhibited in a triangular array, or 'trellis' diagram, gave estimates of the size of the moth population for each day. The interpretation of trellis diagrams was further developed by Fisher & Ford (1947). They estimated daily numbers, taking into account a death-rate which was obtained from the average time interval between marking and recapture. A more detailed application of this procedure was made by Dowdeswell *et al.* (1949).

Although in the various papers cited above effective use has been made of *estimates* of population size and of birth- and death-rates, there has been little discussion for the most part of the *precision* of the results obtained. Jackson (1937, 1939) obtained both the population estimate and its variance for the 'negative' and 'positive' methods by fitting a curve

to a series of 'standardized' recapture rates. However, this was not really adequate, as it did not employ satisfactory weighting factors. Schnabel (1938) considered the problem of finding the total number of fish in a lake, and obtained the maximum-likelihood estimate of the population size for repeated sampling of the recapture type. Although approximations to the solution of the likelihood equations were given, there was no discussion of the precision of the results. De Lury (1947) has used a somewhat different approach to the same problem. Specimens caught are removed from the population, and it is assumed that catches are sufficiently large to produce an appreciably diminishing catch per unit effort with repeated catching. An appendix to his paper gives a full discussion of the appropriate statistical model.

It is the purpose of the present paper to consider the precision of estimates of population size, and of birth- and death-rates derived from recapture data. We shall deal first with simple recapture, considering both direct and inverse sampling, and then go on to a new discussion of Jackson's 'negative' and 'positive' methods. Although, in general, the simultaneous estimation of population size and birth- and death-rates from a large trellis diagram is exceedingly tedious, it will be shown that in at least one special case, that of 'triple capture' relating to catching on only three occasions, a satisfactory explicit solution is available. In certain circumstances, therefore, it might be worth while making a special attempt to collect data in this form.

We shall, in the present discussion, assume that we are dealing with closed regions, unaffected by migration. Further, when birth- and death-rates are involved, it will be assumed that the numbers of marked animals are sufficiently large for these rates to operate deterministically.

2. Simple recapture

(*a*) *Direct sampling*

Suppose that we capture, mark and release a animals out of a total population of size x. When the marked animals have freely mingled with the unmarked, we catch a random sample of size n, of which r are found to be marked. We shall assume that n is sufficiently small compared with x for us to be able to ignore the complications of sampling without replacement. The large sample theory of the maximum-likelihood approach will be appropriate if we envisage n, a and x all tending to infinity while keeping constant ratios to each other. The likelihood is

$$e^L = \binom{n}{r}\left(\frac{a}{x}\right)^r\left(\frac{x-a}{x}\right)^{n-r}. \tag{2·1}$$

Therefore

$$L = \text{const.} + (n-r)\log(x-a) - n\log x. \tag{2·2}$$

Although x is in fact discrete, it will be convenient to treat it as though it were a continuous parameter. Differentiating with respect to x gives

$$\frac{\partial L}{\partial x} = \frac{n-r}{x-a} - \frac{n}{x}. \tag{2·3}$$

Hence the maximum-likelihood estimate of x is

$$\hat{x} = an/r. \tag{2·4}$$

Schnabel (1938) discussed the solution of the corresponding likelihood equation for a set of independent samples, e.g.

$$\sum_i (n_i - r_i)/(x - a_i) = \sum_i n_i/x,$$

where the suffix i indicates the values of n, r and a appropriate to the ith sample. She did not, however, go on to examine the precision of the estimate thus obtained.

Differentiating (2·3) with respect to x, and then taking the expectation provides the required information in the case of a single sample. Thus

$$I_x = -E\frac{\partial^2 L}{\partial x^2} = an/\{x^2(x-a)\}. \tag{2·5}$$

Using the value of x given by (2·4) we can then write

$$\operatorname{var}\hat{x} = I_x^{-1} = a^2 n(n-r)/r^3. \tag{2·6}$$

Now $\hat{x}$ has an infinite expectation if we admit the value $r = 0$. On the other hand, this value will occur so rarely in large samples, for which $m = Er$ is not small, that we can choose to exclude it. With this convention it can be shown that the relative bias in the expectation of $\hat{x}$ is of order m^{-1}. For, adopting the technique employed by Anscombe (1948), we can expand the right-hand side of (2·4) as a Taylor series in powers of $t = r - m$. Taking expectations of both sides then leads to an asymptotic formula for $E\hat{x}$ $(r \neq 0)$. We find

$$\underset{r\neq 0}{E}\,\hat{x} \sim x\left(1+\frac{x-a}{an}+O(m^{-2})\right) = x\left(1+\frac{1}{m}+O(m^{-2})\right), \tag{2·7}$$

if m is small compared with n.

Suppose $x = 1000$, and $n = a = 100$. Then $m = 10$, and although the sample size is as large as 100, there is an average relative bias of the order of 10 % in the maximum-likelihood estimate of x. A slightly adjusted estimate which suggests itself is

$$\check{x} = a(n+1)/(r+1). \tag{2·8}$$

The expectation of $\check{x}$ can be found exactly. We have

$$\begin{aligned} E\check{x} &= \sum_{r=0}^{n} \frac{a(n+1)}{(r+1)}\binom{n}{r}\left(\frac{a}{x}\right)^r\left(\frac{x-a}{x}\right)^{n-r} \\ &= x\sum_{r=0}^{n}\binom{n+1}{r+1}\left(\frac{a}{x}\right)^{r+1}\left(\frac{x-a}{x}\right)^{n-r}. \end{aligned}$$

Therefore

$$E\check{x} = x\left\{1-\left(\frac{x-a}{x}\right)^{n+1}\right\}. \tag{2·9}$$

For large n

$$\left(\frac{x-a}{x}\right)^{n+1} \sim e^{-m}, \tag{2·10}$$

which is quite small even for moderate m. For the values given above, the relative bias is now less than $2{\cdot}5 \times 10^{-5}$. A satisfactory expression for the variance of $\check{x}$ is somewhat more difficult to find. However, using the expansion technique referred to above we can derive the asymptotic series

$$E\check{x}^2 \sim \frac{a^2(n+1)^2}{m^2}\left(1+\frac{1-2p}{m}+\ldots\right), \tag{2·11}$$

where $p = a/x$. We now obtain the variance from the relation

$$\operatorname{var}\check{x} = E\check{x}^2 - (E\check{x})^2. \tag{2·12}$$

We could obtain a single-series expression for the variance by substituting (2·11) in (2·12), and writing

$$(E\check{x})^2 \sim x^2 = a^2n^2/m^2 = \frac{a^2(n+1)^2}{m^2}\left(1-\frac{2}{n}+\ldots\right),$$

but as this would involve the consideration of inverse powers of n, there seems little to be gained by this device.

Fortunately, there exists a nearly unbiased estimate of the variance of $\check{x}$, which is quite convenient for use with samples that are not too small.

It is easy to show that

$$E\,\frac{a^2(n+1)(n+2)}{(r+1)(r+2)} = x^2\left\{1-\left(\frac{x-a}{x}\right)^{n+2}-(n+2)\left(\frac{a}{x}\right)\left(\frac{x-a}{x}\right)^{n+1}\right\} \tag{2·13}$$

$$\sim x^2(1-m\,e^{-m}), \quad \text{for large } n. \tag{2·14}$$

Therefore, if we write

$$T = \frac{a^2(n+1)^2}{(r+1)^2} - \frac{a^2(n+1)(n+2)}{(r+1)(r+2)} = \frac{a^2(n+1)(n-r)}{(r+1)^2(r+2)}, \tag{2·15}$$

then

$$ET = E\check{x}^2 - x^2(1-m\,e^{-m}) \sim \sigma_{\check{x}}^2 + x^2 m\,e^{-m}, \tag{2·16}$$

neglecting quantities of order e^{-m}. Now we know from (2·5) that $\sigma_{\check{x}}^2$ is of order x^2/m, so that

$$ET \sim \sigma_{\check{x}}^2(1+m^2e^{-m}). \tag{2·17}$$

The relative bias of T in our previous example is thus of order $2{\cdot}5\times10^{-3}$, which should be quite satisfactory for practical purposes.

It is worth mentioning here that in certain ecological problems we may be more concerned to use the reciprocal of the population size, $1/x$, as the appropriate index, rather than the population size itself. It follows from (2·4) that the maximum-likelihood estimate of this reciprocal is r/an. Consideration of the binomial distribution in (2·1) shows that this estimate is unbiased and has variance $(x-a)/(anx^2)$.

(*b*) *Inverse sampling*

Now in the above treatment we studied the mean and variance of estimates of the population size x in relation to repeated sampling for which the total size of the samples drawn was regarded as fixed. Various writers (e.g. Haldane, 1945) have considered the method of *inverse sampling*, according to which random sampling is continued until a certain predetermined number of individuals with a particular attribute has been obtained. The use of inverse sampling with simple recapture results in an appreciable simplification—an unbiased estimate of population size and an exact value of the sampling variance are available, even in the most general case of sampling without replacement. It should, of course, be mentioned in passing that it may not always be possible to adopt this method of sampling in practice.

Suppose we catch, mark and release a animals in a total population of size x. We then draw a random sample by the inverse method, continuing our catching until we have m marked animals altogether. The total number of animals in the sample is n, which is now the random variable under consideration.

The probability of obtaining a sample of size n is the probability of drawing first a sample of size $(n-1)$ containing $(m-1)$ marked animals, followed by the drawing of just one further

marked animal. Having regard to the existence of sampling without replacement, the likelihood is clearly

$$P(n) = \frac{\binom{a}{m-1}\binom{x-a}{n-m}}{\binom{x}{n-1}} \frac{(a-m+1)}{(x-n+1)} \tag{2·18}$$

$$= \frac{a}{x}\binom{a-1}{m-1}\binom{x-a}{n-m}\Big/\binom{x-1}{n-1}, \tag{2·19}$$

where $0 \leqslant m \leqslant a$ and $m \leqslant n \leqslant x+m-a$.

It follows from (2·19) that

$$\sum_{n=m}^{x+m-a}\binom{x-a}{n-m}\Big/\binom{x-1}{n-1} = \frac{x}{a\binom{a-1}{m-1}}\sum_{n=m}^{x+m-a}P(n) = x\Big/\left\{a\binom{a-1}{m-1}\right\}. \tag{2·20}$$

The expectations of n and $n(n+1)$, which we shall have occasion to use below, are readily derived as follows:

$$En = \sum_{n=m}^{x+m-a}\frac{na}{x}\binom{a-1}{m-1}\binom{x-a}{n-m}\Big/\binom{x-1}{n-1}$$

$$= a\binom{a-1}{m-1}\sum_{n=m}^{x+m-a}\binom{x-a}{n-m}\Big/\binom{x}{n}. \tag{2·21}$$

The summation on the right-hand side of (2·21) is obtained immediately from (2·20) on writing $x+1$, $a+1$, $n+1$ and $m+1$ for x, a, n and m respectively. Therefore

$$En = a\binom{n-1}{m-1}(x+1)\Big/(a+1)\binom{a}{m} = \frac{m(x+1)}{(a+1)}. \tag{2·22}$$

Similarly,

$$En(n+1) = \sum_{n=m}^{x+m-a}\frac{n(n+1)a}{x}\binom{a-1}{m-1}\binom{x-a}{n-m}\Big/\binom{x-1}{n-1}$$

$$= (x+1)a\binom{a-1}{m-1}\sum_{n=m}^{x+m-a}\binom{x-a}{n-m}\Big/\binom{x+1}{n+1}. \tag{2·23}$$

To evaluate the summation on the right-hand side of (2·23) we write $x+2$, $a+2$, $n+2$ and $m+2$ for x, a, n and m in equation (2·20). Therefore

$$En(n+1) = (x+1)a\binom{a-1}{m-1}(x+2)\Big/(a+2)\binom{a+1}{m+1} = \frac{m(m+1)(x+1)(x+2)}{(a+1)(a+2)}. \tag{2·24}$$

Now, strictly regarded, x can take only integral values. Thus the greatest value attained by the likelihood is derived by considering the ratio

$$\frac{P(n\,|\,x+1)}{P(n\,|\,x)} = \frac{(x-a+1)(x-n+1)}{(x+1)(x-a-n+m+1)}, \tag{2·25}$$

using (2·19). It follows from (2·25) that

$$\left.\begin{aligned} &P(n\,|\,x+1) \gtrless P(n\,|\,x), \\ &\text{according as} \quad x \lessgtr \frac{an}{m} - 1. \end{aligned}\right\} \tag{2·26}$$

Unless an/m is an integer, the likelihood attains its greatest value when x is equal to the integral part of an/m. If an/m is integral the greatest value is attained at both $\left(\frac{an}{m}-1\right)$ and $\frac{an}{m}$.

In practice x is usually sufficiently large for us to be able to ignore its discreteness. In any case we are led to consider the estimates of the general type an/m. It follows from (2·22) that

$$x' = \frac{n(a+1)}{m} - 1 \tag{2·27}$$

is an unbiased estimate of x, for any values of x, n, a and m whatever. In large samples x' approximates to the maximum-likelihood solution. The exact small sample variance of x' is also easily evaluated, using the results in (2·22) and (2·24). We find

$$\text{var}\, x' = \frac{(a-m+1)(x+1)(x-a)}{m(a+2)}. \tag{2·28}$$

3. Jackson's 'negative' and 'positive' methods

(a) *The 'negative' method*

The 'negative' method was first described by Jackson (1937), and simply consists in the repeated catching, marking and releasing of animals on successive occasions, making no record of the number of recaptures until the last occasion. Under certain circumstances, when the preliminary marking can be carried out by relatively unskilled workers while the final scoring of the numbers of recaptures requires a more elaborate set-up, this procedure may be very suitable.

As mentioned in the introduction above, Jackson's method of fitting a curve to 'standardized' recapture rates is not really satisfactory owing to the lack of proper weighting factors. The present section will develop the maximum likelihood approach.

Suppose that on the jth day previous to the last day a_j freshly marked animals are released. Then, if we can assume a constant death-rate γ, $a_j e^{-\gamma j}$ of these animals will survive to stand a chance of recapture on the last day. Of a total of n captures on the last day, let r_j have been marked on the jth day previously and let r_0 be unmarked. With animals bearing more than one mark it will be convenient to take notice of the earliest mark only, although this will result in the loss of a slight amount of information. Then if x is the size of the population on the last day, we can write the likelihood as

$$e^L \propto \prod_{j=1} \left(\frac{a_j e^{-\gamma j}}{x}\right)^{r_j} \left(\frac{x - \sum_{j=1} a_j e^{-\gamma j}}{x}\right)^{r_0}. \tag{3·1}$$

Therefore

$$L = \text{const.} - n \log x - \gamma \sum_{j=1} j r_j + r_0 \log\left(x - \sum_{j=1} a_j e^{-\gamma j}\right). \tag{3·2}$$

Let us write

$$\left.\begin{aligned} F(\gamma) &= \sum_{j=1} a_j e^{-\gamma j}, \\ F'(\gamma) &= -\sum_{j=1} j a_j e^{-\gamma j}, \\ F''(\gamma) &= \sum_{j=1} j^2 a_j e^{-\gamma j}. \end{aligned}\right\} \tag{3·3}$$

Further, put
$$\sum_{j=1} jr_j = A. \tag{3·4}$$

Then the likelihood equations are

$$\frac{\partial L}{\partial x} = -\frac{n}{x} + \frac{r_0}{x - F(\gamma)} = 0 \tag{3·5}$$

and
$$\frac{\partial L}{\partial \gamma} = -A - \frac{r_0 F'(\gamma)}{x - F(\gamma)} = 0. \tag{3·6}$$

We can eliminate x from (3·5) and (3·6) to give

$$\sum_{j=1} c_j e^{-\hat{\gamma} j} = 0, \tag{3·7}$$

where
$$c_j = a_j\left(j - \frac{A}{n - r_0}\right). \tag{3·8}$$

The quantities c_j are readily determined from the observations, and the required root, $\hat{\gamma}$, of (3·7) is easily obtained in practice by the usual iterative procedure. Substituting this value of γ in (3·5) then gives

$$\hat{x} = nF(\hat{\gamma})/(n - r_0). \tag{3·9}$$

Further differentiation of (3·5) and (3·6), followed by taking expectations, gives the amounts of information

$$\left.\begin{aligned} I_{xx} &= nF/\{x^2(x-F)\}, \\ I_{x\gamma} &= -nF'/\{x(x-F)\}, \\ I_{\gamma\gamma} &= n\{F'^2 + F''(x-F)\}/\{x(x-F)\}, \end{aligned}\right\} \tag{3·10}$$

where F, F' and F'' are calculated for $\gamma = \hat{\gamma}$. It now follows from (3·10) that

$$\operatorname{var} \hat{x} = \frac{x^2}{n}\left(\frac{xF''}{FF'' - F'^2} - 1\right), \tag{3·11}$$

and
$$\operatorname{var} \hat{\gamma} = xF/\{n(FF'' - F'^2)\}. \tag{3·12}$$

It is, of course, important to realize that the foregoing treatment assumes that the death-rate operates 'deterministically'. This will be satisfactory if the numbers of marked animals involved are sufficiently large; otherwise, 'stochastic' fluctuations may become important, and it is possible that, under such conditions, the variances of the estimates would be larger.

Example. Jackson (1939, p. 241) gives some data for the 'negative' method carried out on a tsetse fly population. In our notation we have:

j	6	5	4	3	2	1	0
a_j	1262	1086	1299	1401	1198	1183	—
r_j	1	5	17	28	48	70	1389

The total number, n, captured on the last day ($j = 0$) was 1558. It is not clear what convention was adopted with regard to flies marked more than once, but for the purpose of this illustration we will assume that on the last day notice was taken of the first mark only.

The quantity A in (3·4) is 349, and the six values of c_j are easily found to be

$$c_1 = -1260{\cdot}00, \quad c_4 = 2513{\cdot}45,$$
$$c_2 = -\ 77{\cdot}98, \quad c_5 = 3187{\cdot}31,$$
$$c_3 = \ 1309{\cdot}81, \quad c_6 = 4965{\cdot}86.$$

We now have to solve (3·7). Writing $e^{-\gamma} = \mu$ we take the trial values $\mu = 0{\cdot}55$ and $0{\cdot}54$; the right-hand side of (3·7) is then $+29{\cdot}193$ and $-13{\cdot}692$ respectively. Interpolating gives

$$\left.\begin{aligned} \hat{\mu} &= 0{\cdot}543, \\ \hat{\gamma} &= -\log_e \hat{\mu} = 0{\cdot}611. \end{aligned}\right\}$$

The next stage is to compute the F's from (3·3). We find

$$\left.\begin{aligned} F(\hat{\gamma}) &= \ 1416{\cdot}444, \\ F'(\hat{\gamma}) &= -2923{\cdot}873, \\ F''(\hat{\gamma}) &= \ 8327{\cdot}083. \end{aligned}\right\}$$

We are now in a position to substitute in (3·9), (3·11) and (3·12), obtaining finally the following estimates with attached standard errors:

$$\hat{x} = 13{,}060 \pm 1890,$$
$$\hat{\gamma} = \ 0{\cdot}611 \pm 0{\cdot}060.$$

(b) The 'positive' method

The 'positive' method was also first described by Jackson (1937). This consists in marking a large number of animals on a single occasion. On each of a number of later occasions the proportion of marked animals is recorded. The initial population can then be estimated from these data, taking the birth-rate into account. Jackson again used the method of fitting a curve to 'standardized' recapture rates, as in the 'negative' method, and the same objection applies.

The analysis of this type of data is of necessity more complicated as it entails sampling on a number of occasions, whereas in the 'negative' method sampling variation occurs only on the last day. We shall find that the maximum-likelihood treatment is more involved, but is fairly readily amenable to the usual method of calculating scores and information functions.

In order to render the mathematical treatment as simple as possible it is desirable that nearly all the animals captured on each occasion be released again, otherwise, unless numbers are very large, it would be necessary to take into account the depletion of marked animals by the process of recapture.

If we introduce a death-rate in this case we shall find that the corresponding factors in the likelihood cancel, as we should expect. It is, however, necessary to consider a birth-rate, β. Thus after j days the initial population of size x will have become $x e^{\beta j}$. Suppose that, initially, we release a marked animals, and that j days later we catch n_j animals, r_j of which are marked. It is easy to see that

$$L = \text{const.} - \log x \sum_{j=1} n_j - \beta \sum_{j=1} j n_j + \sum_{j=1} (n_j - r_j) \log (x e^{\beta j} - a). \qquad (3{\cdot}13)$$

The likelihood scores are $$S(x) = \frac{\partial L}{\partial x} = \sum_{j=1} \frac{n_j - r_j}{x - a e^{-\beta j}} - \frac{1}{x} \sum_{j=1} n_j, \tag{3·14}$$

and $$S(\beta) = \frac{\partial L}{\partial \beta} = x \sum_{j=1} \frac{j(n_j - r_j)}{x - a e^{-\beta j}} - \sum_{j=1} j n_j. \tag{3·15}$$

Proceeding in the usual way, we find

$$\left.\begin{aligned} I_{xx} &= \frac{a}{x^2} \sum_{j=1} n_j/(x e^{\beta j} - a), \\ I_{x\beta} &= \frac{a}{x} \sum_{j=1} j n_j/(x e^{\beta j} - a), \\ I_{\beta\beta} &= a \sum_{j=1} j^2 n_j/(x e^{\beta j} - a). \end{aligned}\right\} \tag{3·16}$$

Using (3·14), (3·15) and (3·16), we can follow the normal method of scoring. Thus if θ is a column vector representing approximate values of $\hat{x}$ and $\hat{\beta}$, if S is the corresponding vector of scores, and if I is the information matrix, then improved approximations to the maximum-likelihood estimates are given by θ_1, where

$$\theta_1 = \theta + I^{-1} S.$$

As with the death-rate in the previous section, so here we have assumed that numbers are sufficiently large for us to work with a deterministic birth-rate. With small numbers we should have to consider stochastic fluctuations.

4. TRELLIS DIAGRAMS

Jackson's 'negative' and 'positive' methods, discussed in the previous section, are clearly special cases of a more general procedure in which all, or at least the majority, of the animals caught on a particular day are marked and released; at the same time all captured animals are carefully scored with respect to the days on which they were previously marked. As mentioned in the last section, it is convenient, when an animal has more than one marking, to take notice of the earliest mark only, although this must result in a slight loss of information. The record showing the numbers of captures, releases and recaptures can then be set out in a triangular array. Such arrays have been used with great effect by Dowdeswell *et al.* (1940, 1949) and Fisher & Ford (1947) in the analysis of Lepidoptera populations. An example of such an array for three days only is shown in § 5 of the present paper. Fisher & Ford have analysed these trellis diagrams, taking into account a death-rate estimated from the average time interval between release and recapture. Daily estimates of population size can then be made. While this method is extremely valuable in practice it is difficult to assess the precision of the results obtained.

Let us introduce a general notation for these arrays. On the jth day let the total catch be n_j, of which n_{ji} were first marked on the ith day; while n_{j0} are unmarked. Further, let s_j freshly marked animals be released on the jth day. The total population on the jth day is to be represented by N_j, of which N_{ji} were first marked on the ith day, and N_{j0} are unmarked. Then we can attempt a large sample theory using the method of maximum likelihood, by adopting the following approach.

If we can assume that the N_j and N_{ji} are sufficiently large for us to ignore the effects of sampling without replacement, then the likelihood is evidently given by

$$e^L \propto \prod_{j=2}^{k} \prod_{i=0}^{j-1} (N_{ji}/N_j)^{n_{ji}}, \tag{4·1}$$

where the trellis diagram extends over k days in all. If we can further assume the existence of constant birth- and death-rates, we have

$$\left.\begin{aligned} N_{ji} &= s_i e^{-\gamma(j-i)} \quad (i = 1, \ldots, (j-1)), \\ N_j &= N_1 e^{(\beta-\gamma)(j-1)} \quad (j = 2, \ldots, k), \\ \text{and} \quad N_{j0} &= N_j - \sum_{i=1}^{j-1} N_{ji}. \end{aligned}\right\} \tag{4·2}$$

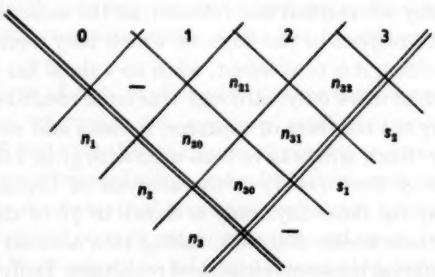
It will be realized that while the whole population is subject to both birth and death, the group of marked individuals suffers depletion by death but gains new recruits only from the release of newly marked specimens.

We can now substitute (4·2) in (4·1) and proceed with the ordinary maximum-likelihood estimation of β, γ and, say, N_1. This is a laborious undertaking in general, and in any case is hardly applicable to data where the numbers of recaptures are low, as with much of the Lepidoptera material, for the reasons discussed in § 2 above, though it might be worth while with the fairly abundant data such as is often available with tsetse fly populations.

However, this approach is susceptible of a fairly simple treatment if we restrict ourselves to catching on three occasions only. This will now be discussed in the following section.

5. The 'triple-catch' method

Suppose that we confine ourselves to catching on only three occasions. Then in the notation of § 4, the numbers caught, released and recaptured are as shown below:



Let us assume, as in the two previous sections, that it is legitimate to work with deterministic birth- and death-rates, which are taken to be β and γ respectively. Thus a group of N individuals gives rise to a group of size $N e^{(\beta-\gamma)t}$ after time t. Note, however, that newly born or emerged individuals are always unmarked, so that the *number* of marked individuals is subject only to the death-rate. Also suppose that only a few captured specimens are not subsequently released again, so that the numbers of animals marked on different days in the population are not appreciably affected.

Finally, let the time intervals between days 1 and 2, and between days 2 and 3, be t_1 and t_2 respectively. Then, in the notation of § 4, the expected numbers of marked and unmarked animals in the population are given by

$$\left.\begin{aligned} N_{21} &= s_1 e^{-\gamma t_1}, & N_{32} &= s_2 e^{-\gamma t_2}, \\ N_{20} &= N_1 e^{(\beta-\gamma)t_1} - s_1 e^{-\gamma t_1}, & N_{31} &= s_1 e^{-\gamma(t_1+t_2)}, \\ N_2 &= N_1 e^{(\beta-\gamma)t_1}, & N_{30} &= N_1 e^{(\beta-\gamma)(t_1+t_2)} - s_2 e^{-\gamma t_2} - s_1 e^{-\gamma(t_1+t_2)}, \\ & & N_3 &= N_1 e^{(\beta-\gamma)(t_1+t_2)}. \end{aligned}\right\} \tag{5·1}$$

The general maximum-likelihood treatment is simplified if we put

$$\left.\begin{aligned} N_2 = N_1 e^{(\beta-\gamma)t_1} &= x, \\ e^{\beta t_2} &= \lambda, \\ e^{-\gamma t_1} &= \mu. \end{aligned}\right\} \tag{5·2}$$

The likelihood can now be written

$$e^L \propto x^{-(n_2+n_3)}(x-\mu s_1)^{n_{20}}(\lambda x-\mu s_1-s_2)^{n_{30}}\lambda^{-n_3}\mu^{(n_{21}+n_{31})}. \tag{5·3}$$

The maximum-likelihood equations are therefore

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= -\frac{n_2+n_3}{x} + \frac{n_{20}}{x-\mu s_1} + \frac{n_{30}\lambda}{\lambda x-\mu s_1-s_2} = 0, \\ \frac{\partial L}{\partial \lambda} &= -\frac{n_3}{\lambda} + \frac{n_{30}x}{\lambda x-\mu s_1-s_2} = 0, \\ \frac{\partial L}{\partial \mu} &= \frac{n_{21}+n_{31}}{\mu} - \frac{n_{20}s_1}{x-\mu s_1} - \frac{n_{30}s_1}{\lambda x-\mu s_1-s_2} = 0. \end{aligned}\right\} \tag{5·4}$$

It is worth observing that we are estimating three parameters, and have just three available degrees of freedom, one from the sample on day 2 and two from the sample on day 3. This situation has been discussed by Bailey (1951), who pointed out that the equations given by setting the expected numbers equal to their expectations provide maximum-likelihood solutions, and are often simpler to solve than the likelihood equations themselves. At the same time it is necessary to check that the likelihood equations, regarded as linear functions of the observations, are linearly independent. The appropriate test, that the Jacobian matrix of the expectations with respect to the parameters to be estimated shall have rank equal to the number of those parameters, is easily seen to be satisfied in the present case. Thus we can either solve (5·4) directly, or we can solve the alternative equations, which, using (5·1) and (5·2), are evidently

$$\left.\begin{aligned} n_{21} &= \frac{N_{21}}{N_2}n_2 = n_2\mu s_1/x, \\ n_{32} &= \frac{N_{32}}{N_3}n_3 = n_3 s_2/\lambda x, \\ n_{31} &= \frac{N_{31}}{N_3}n_3 = n_3\mu s_1/\lambda x, \end{aligned}\right\} \tag{5·5}$$

omitting the other two linearly dependent equations. From (5·5) we have, immediately,

$$\left.\begin{aligned} \hat{x} &= s_2 n_2 n_{31}/n_{21} n_{32}, \\ e^{\beta t_2} = \hat{\lambda} &= n_{21} n_3/n_2 n_{31}, \\ e^{-\gamma t_1} = \hat{\mu} &= s_2 n_{31}/s_1 n_{32}. \end{aligned}\right\} \tag{5·6}$$

The large sample variances of these estimates can be most easily obtained from Fisher's formula for the variance of a function of the observations

$$\operatorname{var} T = \sum_i m_i \left(\frac{\partial T}{\partial a_i}\right)^2 - n\left(\frac{\partial T}{\partial n}\right)^2 \Bigg|_{a_i = m_i}, \tag{5·7}$$

where T is a function of the observations a_i, for which the expectations are m_i, the total sample size being n. In the present case there will be terms like $n(\partial T/\partial n)^2$ corresponding to both n_2 and n_3. Thus we have

$$\left.\begin{aligned} \operatorname{var} \hat{x} &= \hat{x}^2\left(\frac{1}{n_{21}} + \frac{1}{n_{32}} + \frac{1}{n_{31}} - \frac{1}{n_2}\right), \\ \operatorname{var} \hat{\lambda} &= \hat{\lambda}^2\left(\frac{1}{n_{21}} + \frac{1}{n_{31}} - \frac{1}{n_2} - \frac{1}{n_3}\right), \\ \operatorname{var} \hat{\mu} &= \hat{\mu}^2\left(\frac{1}{n_{32}} + \frac{1}{n_{31}}\right). \end{aligned}\right\} \tag{5·8}$$

The large sample variances of $\hat{\beta}$ and $\hat{\gamma}$ are therefore

$$\left.\begin{aligned} \operatorname{var} \hat{\beta} &= \operatorname{var} \hat{\lambda}/(\partial\lambda/\partial\beta)^2 = \left(\frac{1}{n_{21}} + \frac{1}{n_{31}} - \frac{1}{n_2} - \frac{1}{n_3}\right)\Big/ t_2^2, \\ \operatorname{var} \hat{\gamma} &= \operatorname{var} \hat{\mu}/(\partial\mu/\partial\gamma)^2 = \left(\frac{1}{n_{32}} + \frac{1}{n_{31}}\right)\Big/ t_1^2. \end{aligned}\right\} \tag{5·9}$$

The results given in (5·8) can be verified by the alternative procedure of calculating and inverting the information matrix, though this is rather more lengthy.

Having regard to the results of § 2, it is evident that the above formulae require n_{21}, n_{32} and n_{31} all to be fairly large. However, estimates of the type shown in (2·8) will be approximately unbiased for much smaller values. There is therefore some advantage in using the adjusted estimates

$$\left.\begin{aligned} \check{x} &= s_2(n_2+1)\, n_{31}/(n_{21}+1)\,(n_{32}+1), \\ e^{\beta t_2} = \check{\lambda} &= n_{21}(n_3+1)/n_2(n_{31}+1), \\ e^{-\gamma t_1} = \check{\mu} &= s_2 n_{31}/s_1(n_{32}+1). \end{aligned}\right\} \tag{5·10}$$

As in § 2 there is some difficulty in calculating the population values of the variances of estimates given in (5·10), but again it is possible to obtain approximately unbiased *estimates* of the variances of $\check{x}$, $\check{\lambda}$ and $\check{\mu}$. Consider, for example, the variance of $\check{x}$ in (5·10). In a manner similar to that adopted in § 2 for simple recapture, we have

$$E\,\frac{(n_2+1)\,(n_2+2)\,n_{31}(n_{31}-1)}{(n_{21}+1)\,(n_{21}+2)\,(n_{32}+1)\,(n_{32}+2)} = E\,\frac{(n_2+1)\,(n_2+2)}{(n_{21}+1)\,(n_{21}+2)}\,E\,\frac{n_{31}(n_{31}-1)}{(n_{32}+1)\,(n_{32}+2)} \sim \frac{N_2^2}{N_{21}^2}\,\frac{N_{31}^2}{N_{32}^2} = \frac{x^2}{s_2^2}. \tag{5·11}$$

Consider, therefore, the statistic

$$U = \check{x}^2 - \frac{s_2^2(n_2+1)(n_2+2)\,n_{31}(n_{31}-1)}{(n_{21}+1)(n_{21}+2)(n_{32}+1)(n_{32}+2)}. \tag{5·12}$$

Using (5·11) we see that

$$EU \sim E\check{x}^2 - x^2 \sim \operatorname{var}\check{x}. \tag{5·13}$$

Thus U is an approximately unbiased estimate of the variance of $\check{x}$. By comparison with (2·17) it is evident that the relative errors introduced are of order $m^2 e^{-m}$, where m is n_{21} or n_{32}. For $\check{\lambda}$ and $\check{\mu}$, we have similarly

$$V = \check{\lambda}^2 - \frac{n_{21}(n_{21}-1)(n_3+1)(n_3+2)}{n_2(n_2-1)(n_{31}+1)(n_{31}+2)}, \tag{5·14}$$

where

$$EV \sim \operatorname{var}\check{\lambda}, \tag{5·15}$$

and the relative error is of order $n_{31}^2 \exp(-n_{31})$. Also

$$W = \check{\mu}^2 - \frac{s_2^2 n_{31}(n_{31}-1)}{s_1^2(n_{32}+1)(n_{32}+2)}, \tag{5·16}$$

where

$$EW \sim \operatorname{var}\check{\mu}, \tag{5·17}$$

and the relative error is of order $n_{32}^2 \exp(-n_{32})$.

6. SUMMARY

The maximum-likelihood estimate of the population size in simple recapture is discussed and shown to have a relative bias of order m^{-1}, where m is the expected number of recaptures (rejecting cases when there are no recaptures). This bias may be serious even if the total sample size is fairly large. A slight adjustment of the estimate gives a quantity whose relative bias is of order e^{-m}, and is therefore satisfactory for much smaller values of m. It is difficult to find an adequate expression for the population value of the variance of this latter statistic. Fortunately, however, these exists an estimate of this variance, with relative bias of order $m^2 e^{-m}$, which is sufficiently good for many practical purposes.

It is also shown that with *inverse* sampling there is an appreciable simplification—an unbiased estimate of population size and an exact value of the sampling variance are available, and are valid for samples of any size in the most general case of sampling without replacement.

Birth- and death-rates have been taken into account by Jackson (1937, 1939) as exemplified in his 'negative' and 'positive' methods. The present paper discusses the maximum-likelihood treatment of both methods, which does not seem to have been given previously, providing in particular properly weighted estimates of the precisions involved.

The data collected by repeated marking and recapture on several occasions can be exhibited in the form of a triangular array or 'trellis' diagram. This was first done by Dowdeswell *et al.* (1939). A general maximum-likelihood treatment of such data is extremely laborious. If we restrict ourselves to collecting data on three occasions only—the 'triple-catch' method —then explicit solutions for the population size and birth- and death-rates are possible. As with simple recapture approximately unbiased estimates of these quantities and approximately unbiased estimates of the corresponding variances are also available. In view of the

simplicity of these results it seems that, in certain circumstances, it might be advantageous to make a special effort to collect data in just this form—concentrating on three occasions only—rather than producing a large trellis diagram which may be difficult to interpret.

I am indebted to Prof. R. A. Fisher for his interest in this paper and for many stimulating discussions.

REFERENCES

ANSCOMBE, F. J. (1948). The transformation of Poisson, binomial and negative binomial data. *Biometrika*, **35**, 246–54.

BAILEY, NORMAN T. J. (1951). Testing the solubility of maximum likelihood equations in the routine application of scoring methods. *Biometrics*, **7**, (at Press).

DE LURY, D. B. (1947). On the estimation of biological populations. *Biometrics*, **3**, 145–67.

DOWDESWELL, W. H., FISHER, R. A. & FORD, E. B. (1940). The quantitative study of populations in the Lepidoptera. I. *Polyommatus icarus* Rott. *Ann. Eugen., Lond.*, **10**, 123–36.

DOWDESWELL, W. H., FISHER, R. A. & FORD, E. B. (1949). The quantitative study of populations in the Lepidoptera. 2. *Maniola jurtina* L. *Heredity*, **3**, 67–84.

FISHER, R. A. & FORD, E. B. (1947). The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. *Heredity*, **1**, 143–74.

HALDANE, J. B. S. (1945). On a method of estimating frequencies. *Biometrika*, **33**, 222–5.

JACKSON, C. H. N. (1933). On the true density of tsetse flies. *J. Anim. Ecol.* **2**, 204–9.

JACKSON, C. H. N. (1937). Some new methods in the study of *Glossina morsitans*. *Proc. Zool. Soc. Lond.* 1936, pp. 811–96.

JACKSON, C. H. N. (1939). The analysis of an animal population. *J. Anim. Ecol.* **8**, 238–46.

JACKSON, C. H. N. (1940). The analysis of a tsetse fly population. *Ann. Eugen., Lond.*, **10**, 332–69.

LINCOLN, F. C. (1930). Calculating waterfowl abundance on the basis of banding returns. *Circ. U.S. Dep. Agric.* no. 118, May 1930.

SCHNABEL, Z. E. (1938). The estimation of the total fish population of a lake. *Amer. Math. Mon.* **45**, 348–50.

A MATHEMATICAL THEORY OF ANIMAL TRAPPING

By P. A. P. MORAN

Institute of Statistics, University of Oxford

Leslie & Davis (1939) have developed a mathematical theory for the estimation of the total population of an animal in a given region from the records of a series of trappings by simple traps (each of which killed the animal) under the assumptions that there is no immigration or emigration and that the births and deaths from natural causes during the period of trapping can be assumed small. The purpose of the present paper is to provide a theory which in some respects is preferable, and to discuss its assumptions and limitations.

We suppose that there are N animals in a definite region in which trapping takes place on k successive occasions. These will usually take place at equal intervals of time, although this is not essential. If possible the position of the traps should be chosen at random on each occasion, but if the average region occupied by each animal is large compared with the total area divided by the total number of traps, this is less important. On each occasion the same number of traps are used, and we denote the total catch on the ith occasion by $C_i (i = 1, \ldots, k)$. Unless the size N of the initial population is very large, the sequence of C's will tend to diminish because of the exhaustion of the population, and from this decrease we wish to estimate N. If the above conditions hold, and especially if the traps are placed at random on each occasion, it is not unreasonable to suppose as a first approximation that the probability of any particular animal being caught on a particular occasion is independent of the animal and the occasion (ways in which this assumption can break down are discussed later). Let this probability be q. Then the expected number caught on the first occasion will be Nq. Writing $p = 1-q$, the expected number remaining will be Np, so that the expected number caught on the second occasion will be Npq. Carrying on the argument we see that the expected values of $C_1, \ldots, C_k$ are

$$Nq, Nqp, Nqp^2, \ldots, Nqp^{k-1}.$$

Then

$$\log E(C_i) = \log N + (i-1)\log p + \log q,$$

where $E(C_i)$ is the expected number caught on the ith occasion. As this means that $\log E(C_i)$ is a linear function of i, one method of dealing with the problem would be to replace the $E(C_i)$ by the observed values C_i, take logarithms, calculate a weighted regression line on i using an estimated standard error for each $\log C_i$, and then find the values of N and q from this regression line. This method not only requires a lot of calculation but is not strictly valid, since successive values of the observed C_i are not statistically independent. For example, if C_1 is larger than Nq, the expected value of C_2, given C_1, is $(N-C_1)q$, which is less than Nqp. Leslie & Davis's method is somewhat similar to the above, but involves further complications and suffers from the same objections.

The following method is based on maximum-likelihood estimation and is much shorter, as well as providing easily calculable estimates of the standard errors of estimation. We must now assume that the probability of any given animal being caught is independent of the number of other animals caught. This assumption would break down if the expected catch was nearly as large as the number of traps, since each trap can operate only once on each trapping occasion. However, in the example considered later and in most practical

circumstances the assumption is probably well verified. If, then, at any stage there are N animals in the population, the chance of catching m of them will be given by the term in the binomial distribution

$$\binom{N}{m} q^m p^{N-m}.$$

If on the first occasion C_1 are caught, the population at risk during the second trapping is $N-C_1$, and the probability of catching C_2 will be

$$\binom{N-C_1}{C_2} q^{C_2} p^{N-C_1-C_2}.$$

Writing $r_s = \sum_1^s C_i$, we therefore see that the joint probability of catches $C_1, C_2, \ldots, C_k$ is

$$\binom{N}{C_1} q^{C_1} p^{N-C_1} \ldots \binom{N-r_{k-1}}{C_k} q^{C_k} p^{N-r_k} = \frac{N!}{C_1! \ldots C_k! (N-r_k)!} q^{r_k} p^{kN-\sum_1^k r_i}, \tag{1}$$

and this is the likelihood of the observed catches. Writing L for the natural logarithm of this likelihood and using Stirling's approximation in the form

$$\log n \doteqdot (n+\tfrac{1}{2}) \log n - n + \text{constant},$$

we find

$$L = \text{constant} + (N+\tfrac{1}{2}) \log N - N - (N-r_k+\tfrac{1}{2}) \log (N-r_k) + (N-r_k) + r_k \log q + \left(kN - \sum_1^k r_i\right) \log p.$$

To obtain estimators of N and q we maximize L for variations in N and q by differentiating partially and equating the resulting expressions to zero. We then obtain

$$\frac{\partial L}{\partial q} = \frac{r_k}{q} - \frac{kN - \sum_1^k r_i}{1-q} = 0, \tag{2}$$

$$\frac{\partial L}{\partial N} = \tfrac{1}{2}\{N^{-1} - (N-r_k)^{-1}\} - \log\{(N-r_k)/N\} + k \log (1-q) = 0. \tag{3}$$

It will be found in most cases that the term $\frac{1}{2}\{N^{-1} - (N-r_k)^{-1}\}$ is small enough to be neglected. Solving (1) for q we can therefore write (2) and (3) in the forms

$$\hat{q} = \frac{r_k}{k\hat{N} - \sum_1^{k-1} r_i}, \tag{4}$$

and

$$\log\left(1 - \frac{r_k}{\hat{N}}\right) = k \log (1-\hat{q}), \tag{5}$$

where $\hat{N}$ and $\hat{q}$ are estimators of N and q. Equation (5) is, in fact, equivalent to equating the observed total catch r_k to its expected value. Equations (4) and (5) now have to be solved by a process of successive approximation. One way to do this is to make a guess at $\hat{N}$, then calculate $\hat{q}$ from (4), and then a new value of $\hat{N}_1$ from (5). By doing this for several values of $\hat{N}$ and plotting $\hat{N}_1 - \hat{N}$ it is easy to find the value at which they are approximately equal. One method of estimating a suitable value of N for starting this process is to guess from the data the value of i for which C_i is about one-half of C_1, and take twice the value of r_i as a guess at $\hat{N}$. The process of solving (4) and (5) is much quicker than the use of a regression line.

The asymptotic optimum properties of maximum-likelihood estimators in the present problem do not follow from the usual theorems because the sample size is itself a random variable but, as in other similar problems, seem to remain true because, as we notice from (1), r_k and $\sum_1^k r_i$ are jointly sufficient for N and q. To obtain standard errors for our estimates we shall therefore assume these properties to hold and the variance-covariance matrix of $(\hat{N}, \hat{q})$ is then

$$\begin{pmatrix} -\frac{\partial^2 L}{\partial q^2} & -\frac{\partial^2 L}{\partial N \partial q} \\ -\frac{\partial^2 L}{\partial N \partial q} & -\frac{\partial^2 L}{\partial N^2} \end{pmatrix}^{-1}, \tag{6}$$

where

$$\frac{\partial^2 L}{\partial N^2} = \tfrac{1}{2}\{(N-r_k)^{-2} - N^{-2}\} + N^{-1} - (N-r_k)^{-1}, \tag{7}$$

$$\frac{\partial^2 L}{\partial N \partial q} = -\frac{k}{1-q}, \tag{8}$$

$$\frac{\partial^2 L}{\partial q^2} = -\frac{\left(kN - \sum_1^k r_i\right)}{(1-q)^2} - \frac{r_k}{q^2}. \tag{9}$$

In (7) the first term can usually be neglected.

Leslie & Davis discuss an actual example. An area of 22·5 acres in Freetown, Sierra Leone, was trapped on eighteen successive occasions over a period of six weeks. 210 traps were set in houses on each occasion in a manner which was effectively at random. The number, C_i, of *Rattus rattus* captured on each occasion i is given in Table 1. The greatest number of other animals caught on any one occasion was sixteen. These figures show that on no occasion was the total number of animals trapped greater than fifty-nine, so that the assumption of successive binomial distributions seems a reasonable one.

Table 1

Trapping	C_i	r_i	$E(C_i)$
1	49	49	39·32
2	32	81	36·35
3	31	112	33·60
4	34	146	31·06
5	16	162	28·71
6	33	195	26·54
7	22	217	24·53
8	27	244	22·68
9	17	261	20·96
10	19	280	19·38
11	18	298	17·91
12	16	314	16·56
13	18	332	15·31
14	12	344	14·15
15	14	358	13·08
16	12	370	12·09
17	17	387	11·17
18	7	394	10·33

From the figures of this table, the equations we have to solve are

$$\hat{q} = \frac{394}{18\hat{N} - 4150}, \tag{10}$$

and

$$\log\left(1 - \frac{394}{\hat{N}}\right) = 18\log(1-\hat{q}). \tag{11}$$

We have omitted the term $\frac{1}{2}\{N^{-1} - (N - r_k)^{-1}\}$ from equation (11), as it can be shown to have only a small effect on the solution. We can then, for convenience, take the logarithms to the base 10.

To solve (10) and (11) a method of successive approximation was used. Trial values 500 and 600 were taken for $\hat{N}$, $\hat{p}$ was found from (10) and a new value of $\hat{N}$ found from (11). The difference between the new value of $\hat{N}$ and the old was plotted graphically against the old value to give a better guess. A further approximation gave $N = 520, q = 0{\cdot}0756$, which was considered sufficiently accurate. Inserting these values in (6) gave

$$\begin{pmatrix} 74530 & 19{\cdot}47 \\ 19{\cdot}47 & 0{\cdot}006013 \end{pmatrix}^{-1} = \begin{pmatrix} 0{\cdot}00008707 & 0{\cdot}2819 \\ 0{\cdot}2819 & 1079 \end{pmatrix},$$

from which it follows that

$$\text{S.E.}\,(\hat{N}) = 32{\cdot}9, \quad \text{S.E.}\,(\hat{q}) = 0{\cdot}00933,$$

and confidence limits for $\hat{N}$ and $\hat{q}$ were therefore judged to be 520 ± 66 and $0{\cdot}0756 \pm 0{\cdot}0187$ respectively. Leslie & Davis in their paper found 485 ± 74, but their calculation involved finding an unweighted regression line. Mr P. H. Leslie has kindly communicated to me a recalculation of his data using the estimated weights given in his paper, and this gives $N = 504$ with a standard error of 36.

The expected values of C_i, using $N = 520$, $q = 0{\cdot}0756$, are also shown in Table 1, and comparison with the observed values gives a χ^2 of 17·42 with sixteen degrees of freedom, since two constants have been fitted. Leslie & Davis also calculated a χ^2 from their expected values and found 16·79, which is smaller than the above. However, their method involved estimating another constant derived from the observed number of sprung traps not capturing *R. rattus*. Moreover, maximum-likelihood and minimum χ^2 estimators, although both fully efficient, do not necessarily give the same result. The closeness of these values of χ^2 to their expected values shows that the data are not inconsistent with the hypotheses of either theory.

If the theoretical assumptions underlying the above theory are not true the conclusions may easily become quite incorrect. This may happen in a number of ways, of which the following four are perhaps the most important.

First, if the animals become trap-shy as a result of trapping the theory breaks down and the true population size is then probably greater than that estimated. There is evidence that trap-shyness does occur in some animals, and in the present data there is perhaps some slight indication of it in the rather large drop in the second catch relative to the first.

Secondly, if there exists a positive immigration into the area formula (1) becomes incorrect. The effect of immigration is to make $E(C_i)$ decrease to a constant positive value instead of to zero. If we attempt to take account of this in equation (1) by supposing that the population is increased by a constant number M, say, between each trapping, the estimation problem becomes very complicated because the factorials no longer cancel.

Thirdly, if the population density or liability to capture is so large that most or all of the traps catch animals on each trapping occasion, the theory breaks down. We can no longer assume the number caught to be distributed in the binomial distribution because it is bounded above by the number of traps. The effect on the observations is to make the initial catches nearly constant and equal or nearly equal to the number of traps for as long a period as is necessary to reduce the total population to such a level that the catch is small enough relative to the number of traps to make the binomial distribution a reasonable approximation to the truth. The setting up of a theory to take this into account might be possible but would not be simple. Leslie & Davis do make the number of traps enter into their theory, but not in such a way as to avoid this difficulty. They suggest that it might be overcome by setting a number of traps close together at each trapping point.

The last and perhaps the most important reservation about the above theory is that it assumes that the chance of being trapped is the same for each animal. That this assumption is vital may be seen by imagining an animal population part of which satisfied the assumptions of the theory and part of which was completely trap-shy. The observations would behave exactly as if the assumptions were true, and the estimate obtained would be an estimate only of the number which were not trap-shy, the number of the rest being unknown. Moreover, no test on the observed figures would show the existence of the latter part of the population. In any actual population it would, of course, be more reasonable to assume that there was a distribution of probability of being caught among the individual animals. Unlike the two previous difficulties, it does not seem possible to set up any statistical method of verifying, from the observed catches alone, the correctness of the assumption of equiprobability of being caught.

Finally, it s hould be pointed out that most biologists now use live trapping, marking release and recapture for the estimation of population.

I am very much indebted to Mr P. H. Leslie for his help and criticism during the preparation of this paper, and to Dr Dennis Chitty for some helpful criticism.

REFERENCE

Leslie, P. H. & Davis, D. H. S. (1939). *J. Anim. Ecol.* **8**, 94–113.

SOME SYSTEMATIC EXPERIMENTAL DESIGNS

By D. R. COX

Statistical Laboratory, University of Cambridge

1. Introduction

Consider the following experiment. It is required to compare the effect in processing of a number of treatments applied to wool. The wool is divided into lots as alike as possible and the lots are numbered in random order. In week 1 lot 1 is processed with a certain treatment. In week 2 lot 2 is processed with, in general, a different treatment. And so on. The age of the wool affects its behaviour in processing so that, superimposed on any treatment differences, there will be a smooth trend due to aging. We want to estimate the treatment differences as simply and accurately as possible. In what order should the treatments be applied?

More generally, consider an experiment to compare a number of *treatments* on *plots* arranged in order in space or time, one treatment being applied to each plot. We assume that the measured quantity is

$$\text{(Treatment constant)} + \text{(Smooth trend)} + \text{(Random quantity)}, \qquad (1)$$

where the random quantity has zero mean, constant variance and is independent in different plots. We want to arrange the treatments to give simple and accurate estimates of the treatment differences.

One method is to arrange the experiment in randomized blocks, thus eliminating part of the trend from the error of the treatment comparisons.

A second method is to randomize completely the order of application of the treatments and to eliminate the trend by analysis of covariance on position. The calculations can be extensive if high-order trends are involved. Also when the above assumptions hold information is lost by the use of the method. The loss may be serious if the number of treatments and number of replicates are small (see §4·2).

A third and related method is to repeat the treatments always in the same order, e.g. $T_1T_2T_3T_1T_2T_3\ldots$, and to eliminate the trend by analysis of covariance on position (i.e. by fitting polynomials). Neyman (1929) discussed this in connexion with agricultural field trials and gave certain tables. The calculations are again lengthy, and to simplify them Hald (1948) has introduced generalized orthogonal polynomials the tabulation* of which would greatly reduce the labour of fitting the successive polynomials.

The discussion of a recent paper by Miss G. M. Cox (Cox, 1950) contains some comments on the use of polynomial fitting in the analysis of agricultural field trials. In particular, R. A. Fisher referred to his and Prof. van Uven's early work on the subject and criticized the method.

R. M. Williams (1949) has worked on a related problem in which there is no trend, but the errors in adjacent plots are correlated.

* Hald gives tables in some special cases.

2. Outline of method

Let there be t treatments, n replicates of each treatment and $N = tn$ plots in all. Let the plots be equally spaced and suppose that the trend can be represented by a pth order polynomial

$$\alpha_1 \xi'_1 + \ldots + \alpha_p \xi'_p, \tag{2}$$

where ξ'_μ is the orthogonal polynomial of order μ for N equally spaced points given in Fisher & Yates's tables (1948). To begin with the value of p is assumed known *a priori*. The population mean of the measurement on the jth plot is

$$a_i + \sum_{\mu=1}^{p} \alpha_\mu \xi'_\mu, \tag{3}$$

where a_i depends only on the treatment T_i applied to the jth plot.

The estimation of the a_i is easiest if the treatment differences are *completely orthogonal* to the trend. Now treatment differences are orthogonal to the linear part of the trend if and only if

$$s'_{i1} \equiv \Sigma_i \xi'_1 = 0 \quad (i = 1, \ldots, t), \tag{4}$$

where Σ_i denotes summation over all plots receiving treatment T_i. In general treatment differences are orthogonal to the complete trend if and only if

$$s'_{i\mu} \equiv \Sigma_i \xi'_\mu = 0 \quad (i = 1, \ldots, t;\ \mu = 1, \ldots, p). \tag{5}$$

If (5) is true we may estimate the treatment differences as if the trend were absent. The variance of a treatment comparison is $2\sigma^2/n$, where σ^2 is the error variance. It is shown in §4 that if (5) is not true the variance is in general greater than $2\sigma^2/n$.

Therefore the problem is solved if we can arrange the treatments to satisfy (5) or, if this is impossible, to satisfy (5) as nearly as possible.

We shall say that an arrangement of treatments such that (5) is true forms a design orthogonal to trends of order p.

3. Some exact solutions

When there are only two treatments and the trend can be represented by a cubic, some exact solutions are possible. The total number of plots, $N = 2n$, is even so that all the ξ'_1 are odd integers. Now the sum of an odd number of odd integers cannot be zero, so that it is impossible to satisfy (4) for odd n. Therefore the number of replicates, n, must be even and equal to $2m$, say. Then ξ'_1 takes the form

$$\ldots \quad -5 \quad -3 \quad -1 \quad | \quad 1 \quad 3 \quad 5 \quad \ldots.$$

There exist *symmetrical designs* in which if treatment T_i is applied in position $\xi'_1 = r$ (called position r for short), the same treatment T_i is applied in position $-r$.

Example. $t = 2$, $n = 4$, $N = 8$:

-7	-5	-3	-1	1	3	5	7
T_2	T_2	T_1	T_1	T_1	T_1	T_2	T_2

Since the odd-order orthogonal polynomials $\xi'_{2\beta+1}$ $(\beta = 0, 1, \ldots)$ are odd functions, all symmetrical designs have treatment differences orthogonal to $\xi'_{2\beta+1}$.

The problem is therefore solved if we can find m positions from 1, 3, 5, ... to take treatment T_1, such that

$$\Sigma_1' \xi_2' = 0, \tag{6}$$

where Σ_1' denotes summation over those positions in 1, 3, 5, ... which receive treatment T_1.

There are no solutions for odd m, but there are the following solutions for even m:

Eight replicates. $t = 2, m = 4, n = 8, N = 16, p = 3$:

ξ_1'	1	3	5	7	9	11	13	15
	T_1	T_2	T_2	T_1	T_2	T_1	T_1	T_2

(and the design formed by interchanging T_1 and T_2).

Twelve replicates. $t = 2, m = 6, n = 12, N = 24, p = 3$:

ξ_1'	1	3	5	7	9	11	13	15	17	19	21	23
	T_2	T_2	T_2	T_1	T_1	T_1	T_1	T_2	T_1	T_2	T_1	T_2
	T_2	T_2	T_1	T_2	T_1	T_1	T_1	T_1	T_2	T_2	T_2	T_1
	T_2	T_1	T_2	T_1	T_2	T_1	T_1	T_2	T_2	T_1	T_1	T_2
	T_2	T_2	T_1	T_1	T_2	T_1	T_2	T_1	T_1	T_2	T_1	T_2
	T_2	T_1	T_2	T_1	T_1	T_1	T_2	T_2	T_2	T_1	T_2	T_1
	T_2	T_2	T_1	T_2	T_1	T_2	T_1	T_1	T_1	T_1	T_2	T_2

(and the designs formed by interchanging T_1 and T_2).

Of these the second is recommended for practical use because it is most nearly orthogonal to ξ_4'.

There are probably solutions for all higher even m.

There are also non-symmetric designs orthogonal to ξ_1' and ξ_2'. Examples are:

Four replicates. $t = 2, m = 2, n = 4, N = 8, p = 2$:

ξ_1'	-7	-5	-3	-1	1	3	5	7
	T_2	T_1	T_1	T_2	T_1	T_2	T_2	T_1

Six replicates. $t = 2, m = 3, n = 6, N = 12, p = 2$:

ξ_1'	-11	-9	-7	-5	-3	-1	1	3	5	7	9	11
	T_1	T_2	T_1	T_2	T_2	T_2	T_1	T_1	T_1	T_2	T_1	T_2

Similar solutions exist for $m = 4$, etc., but are of no practical value, since it is better to use the above symmetric designs.

To summarize, there are exact solutions for two treatments as follows:

Cubic trend	Quadratic trend
$n = 8$	$n = 4$
$n = 12$	$n = 6$, etc.
probably $n = 4r$	

A similar analysis for three and four treatments shows that there are no exact solutions for reasonably small n. Therefore it is necessary to consider designs which satisfy (5) only approximately.

4. General properties of approximate solutions*

4·1. *Solution by least squares*

When the treatment effects and the trend are not orthogonal, we set up least-squares equations for the estimates $\hat{a}_1, \ldots, \hat{a}_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$. To do this it is convenient to introduce *normalized* orthogonal polynomials ξ_μ such that

$$\Sigma\xi_\mu = 0, \quad \Sigma\xi_\mu\xi_\nu = \delta_{\mu\nu} = \begin{cases} 1 & (\mu = \nu), \\ 0 & (\mu \neq \nu), \end{cases} \tag{7}$$

where the summation is over all plots. (The ξ_μ are simple multiples of the ξ'_u tabulated in Fisher & Yates.) We write the trend as $\sum_{\mu=1}^{p} \alpha_\mu \xi_\mu$, and the sum of squares to be minimized is then

$$\sum_{i=1}^{t} \Sigma_i \left(x - a_i - \sum_{\mu=1}^{p} \alpha_\mu \xi_\mu \right)^2. \tag{8}$$

The least-squares equations are, after the use of (7),

$$\mathbf{Le} = \mathbf{b}, \tag{9}$$

where $\mathbf{L}$ is the $(t+p)$th order square matrix

$$\left(\begin{array}{c|c} n\mathbf{I}_t & \mathbf{S} \\ \hline \mathbf{S}' & \mathbf{I}_p \end{array} \right), \tag{10}$$

$\mathbf{e}$ is the column vector $\{\hat{a}_1, \ldots, \hat{a}_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_p\}$ and $\mathbf{b}$ is the column vector

$$\{\Sigma_1 x, \ldots, \Sigma_t x, \Sigma\xi_1 x, \ldots, \Sigma\xi_p x\}.$$

In (10) $\mathbf{I}_j$ is the jth order unit matrix, $\mathbf{S}$ is the $t \times p$ matrix whose (i, μ)th element is $\Sigma_i \xi_\mu$, and $\mathbf{S}'$ is the transpose of $\mathbf{S}$. (The general form of the equations can be seen from the special case (13) written out at length below.)

The inverse of the matrix (10) gives the variances and covariances of the $\hat{a}_i$ and hence the variances of the treatment comparisons $(\hat{a}_i - \hat{a}_j)$. Before considering the general case it is best to examine two important special cases.

4·2. *Two treatments*

When there are only two treatments the matrix (10) simplifies considerably. Since $\Sigma\xi_\mu = 0$, $\Sigma_1\xi_\mu + \Sigma_2\xi_\mu = 0$, so that $s_{1\mu} = -s_{2\mu} = s_\mu$, say. Thus

$$\mathbf{L} = \begin{pmatrix} n & 0 & s_1 & \ldots & s_p \\ 0 & n & -s_1 & \ldots & -s_p \\ s_1 & -s_1 & 1 & & \\ \vdots & \vdots & & \ddots & \\ s_p & -s_p & & & 1 \end{pmatrix}.$$

By a Laplace development down the first two columns we get

$$|\mathbf{L}| = n^2 - 2n \sum_{\mu=1}^{p} s_\mu^2.$$

* § 4·5 summarizes the main conclusions of § 4.

Similarly, the cofactors determining $\text{var}(\hat{a}_1)$, $\text{var}(\hat{a}_2)$, $\text{cov}(\hat{a}_1, \hat{a}_2)$ can be found and we get

$$\text{var}(\hat{a}_1 - \hat{a}_2) = \frac{2\sigma^2}{n}\left\{1 - \frac{2}{n}\sum_{\mu=1}^{p} s_\mu^2\right\}^{-1}. \tag{11}$$

Thus the minimum variance, achieved when all s_μ are zero, is the variance of the difference of two independent means. The efficiency E is

$$E = 1 - \frac{2}{n}\sum_{\mu=1}^{p} s_\mu^2. \tag{12}$$

Example. Consider three designs for two treatments, four replicates and cubic trend ($t = 2, n = 4, N = 8, p = 3$).

ξ_1'	ξ_2'	ξ_3'	Order selected at random	Exact solution for quadratic trend	Best solution orthogonal to ξ_1' and ξ_3'
−7	7	−7	T_2	T_2	T_1
−5	1	5	T_1	T_1	T_2
−3	−3	7	T_1	T_1	T_2
−1	−5	3	T_2	T_2	T_1
1	−5	−3	T_2	T_1	T_1
3	−3	−7	T_2	T_2	T_2
5	1	−5	T_1	T_2	T_2
7	7	7	T_1	T_1	T_1
		$\sum_1 \xi_1'$	4	0	0
		$\sum_1 \xi_2'$	6	0	4
		$\sum_1 \xi_3'$	14	16	0
		s_1^2	0·0952	0	0
		s_2^2	0·2143	0	0·0952
		s_3^2	0·7424	0·9697	0
		E	0·4740	0·5152	0·9524

(For example $s_1 = \Sigma_1 \xi_1$, $s_1^2 = (\Sigma_1 \xi_1')^2/168$.)

From this example we conclude that

(i) an order selected at random may be very inefficient;

(ii) the exact solution for $p = 2$ is very inefficient for $p = 3$.

4·3. *Treatments orthogonal to all ξ_μ except one*

Another important special case is obtained when the treatments are orthogonal to all ξ_μ except one, which we call ξ. The matrix $\mathbf{S}$ consists of zeros except in one column which we denote $\{s_1, \ldots, s_l\}$. It is worth writing out the least-squares equations involving the $\hat{a}_i$ in non-matrix form. They are

$$\left.\begin{aligned} n\hat{a}_1 \qquad\qquad + s_1\hat{\alpha} &= \Sigma_1 x, \\ n\hat{a}_2 \qquad\qquad + s_2\hat{\alpha} &= \Sigma_2 x, \\ \vdots \qquad\qquad\quad \vdots & \\ n\hat{a}_l + s_l\hat{\alpha} &= \Sigma_l x, \\ s_1\hat{a}_1 + s_2\hat{a}_2 + \ldots + s_l\hat{a}_l + \hat{\alpha} &= \Sigma\xi x. \end{aligned}\right\} \tag{13}$$

The matrix $\mathbf{L}$ is a simple bordered matrix and

$$|\mathbf{L}| = n^t - n^{t-1}\sum_{i=1}^{t} s_i^2. \tag{14}$$

Similarly, the cofactor of the first element is

$$n^{t-1} - n^{t-2}\sum_{i=2}^{t} s_i^2,$$

so that

$$\text{var}\,(\hat{a}_1) = \frac{\sigma^2\left\{n^{t-1} - n^{t-2}\sum_{i=2}^{t} s_i^2\right\}}{\left\{n^t - n^{t-1}\sum_{i=1}^{t} s_i^2\right\}}. \tag{15}$$

Similarly, we can find $\text{var}\,(\hat{a}_2)$ and $\text{cov}\,(\hat{a}_1, \hat{a}_2)$ and we get

$$\text{var}\,(\hat{a}_1 - \hat{a}_2) = \sigma^2\left\{\frac{2n^{t-1} - n^{t-2}\left(s_1^2 + s_2^2 + 2\sum_{3}^{t} s_i^2\right) - 2n^{t-2}s_1 s_2}{n^t - n^{t-1}\sum_{1}^{t} s_i^2}\right\}. \tag{16}$$

One important consequence is that

$$\text{var}\,(\hat{a}_1 - \hat{a}_2) \geqslant \frac{2\sigma^2}{n} \tag{17}$$

with equality if and only if $s_1 = s_2$.

We can find the variance of any treatment comparison in the same way. One natural measure of the efficiency of the design is the average variance $\overline{V}$ over all possible comparisons of pairs. Any particular $\hat{a}_i$ occurs in $(t-1)$ of the comparisons and does not occur in $(t-1)(t-2)/2$ of them. Therefore

$$\overline{V} = \frac{\sigma^2}{\left\{n^t - n^{t-1}\sum_{1}^{t} s_i^2\right\}}\left\{2n^{t-1} - \frac{2n^{t-2}}{t(t-1)}[(t-1) + (t-1)(t-2)]\sum_{1}^{t} s_i^2 - \frac{4n^{t-2}}{t(t-1)}\sum_{i>j} s_i s_j\right\}.$$

Now $\sum_{1}^{t} s_i = 0$, so that $2\sum_{i>j} s_i s_j = -\Sigma s_i^2$. Thus

$$\overline{V} = \frac{2\sigma^2}{n}\left\{\frac{n - (t-2)/(t-1)\,\Sigma s_i^2}{n - \Sigma s_i^2}\right\},$$

and the average efficiency E is

$$E = \{1 - \Sigma s_i^2/n\}\left\{1 - \frac{(t-2)}{n(t-1)}\Sigma s_i^2\right\}^{-1}. \tag{18}$$

This reduces to a special form of (12) in the special case $t = 2$.

4·4. *General case*

We now return to the general case $t \neq 2$ and $p \neq 1$ (equation (10)). By a Laplace development of (10) on the first t columns we get

$$|\mathbf{L}| = n^t + \sum_{k=1}^{\{p,t\}} (-1)^k n^{t-k}[\Sigma\,|\mathbf{S}^{(k)}|^2], \tag{19}$$

where $\{p, t\}$ is the smaller of p and t and $|\mathbf{S}^{(k)}|$ is the determinant of a kth order minor of the matrix $\mathbf{S}$, the summation being over all possible kth order minors, i.e. we can express $|\mathbf{L}|$ in terms of the sums of squares of all minors that can be formed from $\mathbf{S}$. The cofactor of the first element of $\mathbf{L}$ can be expressed in exactly the same way in terms of the minors of the

matrix $\mathbf{S}_{(1)}$ formed by deleting the first row of $\mathbf{S}$. This leads directly to a formula for var $(\hat{a}_1)$. Similarly, cov $(\hat{a}_1, \hat{a}_2)$ can be obtained in terms of sums of products of certain minors formed from $\mathbf{S}$.

Now in practice we are interested in designs which are nearly orthogonal, i.e. for which the elements of $\mathbf{S}$ are small. In this case, if p and t are not too large, products of more than two of the elements of $\mathbf{S}$ can be neglected, and

$$\left.\begin{aligned} |\mathbf{L}| &\simeq n^t - n^{t-1} \sum_{i,\mu} s_{i\mu}^2, \\ \operatorname{var}(\hat{a}_1) &\simeq \left\{ \frac{n^{t-1} - n^{t-2} \sum\limits_{i+1;\ \mu} s_{i\mu}^2}{|\mathbf{L}|} \right\} \sigma^2, \\ \operatorname{var}(\hat{a}_1 - \hat{a}_2) &\simeq \left\{ \frac{2n^{t-1} - n^{t-2} \sum\limits_{\mu} (s_{1\mu}^2 + s_{2\mu}^2 + 2 s_{1\mu} s_{2\mu} + 2 \sum\limits_{i+1,\ 2} s_{i\mu}^2)}{|\mathbf{L}|} \right\} \sigma^2. \end{aligned}\right\} \tag{20}$$

Averaging over all comparisons of pairs, we get

$$\bar{V} \simeq \frac{2\sigma^2}{n} \left(1 - \frac{(t-2)}{n(t-1)} \sum_{i,\mu} s_{i\mu}^2\right) \left(1 - \frac{1}{n} \sum_{i,\mu} s_{i\mu}^2\right)^{-1},$$

so that the efficiency E is

$$E \simeq \left(1 - \frac{\Sigma s_{i\mu}^2}{n}\right) \left(1 - \frac{(t-2)}{n(t-1)} \Sigma s_{i\mu}^2\right)^{-1}. \tag{21}$$

4·5. *Summary of* §4

The conclusions of §4 about non-orthogonal designs can be summarized as follows:

(i) Least squares equations (9) can be formed for the treatment and trend coefficients.

(ii) From these equations the variance of any treatment comparison can be obtained.

(iii) The solution is simple in the special case of two treatments (§4·2) or when the treatments are orthogonal to all relevant ξ'_μ except one (§4·3).

(iv) The efficiency of the design can be measured in terms of the variance of the difference between two treatments averaged over all comparisons of pairs of treatments. The efficiency is given by (12) and (18) in the above two special cases.

(v) The efficiency in the general case is given by a complicated expression which simplifies to (21) for designs that are nearly orthogonal.

(vi) The greatest accuracy in estimating treatment differences is obtained for orthogonal designs.

5. Some approximate solutions

In this section the application of the results of §4 is illustrated. Consider experiments with three and four treatments and suppose the trend can be represented by a cubic curve ($p = 3$). If we restrict ourselves to symmetric designs automatically orthogonal to ξ'_1 and ξ'_3, the results of §4·3 apply.

In recording symmetric designs, marked by vertical dotted lines, only the treatments for positive ξ'_1 are shown. For example, the first design is in full:

ξ'_1	-11	-9	-7	-5	-3	-1	1	3	5	7	9	11
	T_1	T_2	T_3	T_3	T_2	T_1	T_1	T_2	T_3	T_3	T_2	T_1

In all cases further designs can be formed by interchanging $T_1, T_2, \ldots$. The best designs are:

Three treatments, four replicates ($t = 3, n = 4, N = 12, p = 3$):

ξ_1'	1	3	5	7	9	11	
ξ_2'	−35	−29	−17	1	25	55	$\Sigma\xi_2'^2 = 12{,}012$
	T_1	T_2	T_3	T_3	T_2	T_1	

$$\Sigma_1\xi_2' = 40, \quad \Sigma_2\xi_2' = -8, \quad \Sigma_3\xi_2' = -32,$$
$$\Sigma s_i^2 = \tfrac{2688}{12012} = 0{\cdot}2238,$$
$$\text{Efficiency} = 0{\cdot}9712.$$

Three treatments, six replicates ($t = 3, n = 6, N = 18, p = 3$):

ξ_1'	1	3	5	7	9	11	13	15	17
ξ_2'	−40	−37	−31	−22	−10	5	23	44	68
(i)	T_1	T_2	T_1	T_3	T_2	T_3	T_3	T_2	T_1
(ii)	T_1	T_2	T_3	T_1	T_2	T_3	T_3	T_2	T_1

	(i)	(ii)
$\Sigma_1\xi_2'$	−6	12,
$\Sigma_2\xi_2'$	−6	−6,
$\Sigma_3\xi_2'$	12	−6.

$$\text{Efficiency} = 0{\cdot}9992,$$

Design (ii) is recommended for practical use because it is the more nearly orthogonal to ξ_4'.

Four treatments, four replicates ($t = 4, n = 4, N = 16, p = 3$):

ξ_1'	1	3	5	7	9	11	13	15
ξ_2'	−21	−19	−15	−9	−1	9	21	35
	T_1	T_2	T_3	T_4	T_4	T_3	T_2	T_1

$$\Sigma_1\xi_2' = 28, \quad \Sigma_2\xi_2' = 4, \quad \Sigma_3\xi_2' = -12, \quad \Sigma_4\xi_2' = -20, \quad \text{Efficiency} = 0{\cdot}9796.$$

Four treatments, six replicates ($t = 4, n = 6, N = 24, p = 3$):

ξ_1'	1	3	5	7	9	11	13	15	17	19	21	23
ξ_2'	−143	−137	−125	−107	−83	−53	−17	25	73	127	187	253
	T_1	T_2	T_2	T_3	T_3	T_4	T_4	T_1	T_4	T_1	T_3	T_2

$$\Sigma_1\xi_2' = 18, \quad \Sigma_2\xi_2' = -18, \quad \Sigma_3\xi_2' = -6, \quad \Sigma_4\xi_2' = 6, \quad \text{Efficiency} = 0{\cdot}9999.$$

The following designs are also useful:

Three treatments, three replicates ($t = 3, n = 3, N = 9, p = 2$):

ξ_1'	−4	−3	−2	−1	0	1	2	3	4
ξ_2'	28	7	−8	−17	−20	−17	−8	7	28
(*a*)	T_3	T_2	T_1	T_2	T_1	T_3	T_1	T_3	T_2
(*b*)	T_2	T_3	T_1	T_2	T_1	T_3	T_2	T_3	T_1

(*a*) $\Sigma_1\xi_2' = -36, \quad \Sigma_2\xi_2' = 18, \quad \Sigma_3\xi_2' = 18.$

$$\text{Efficiency} = 0{\cdot}8676.$$

Exactly orthogonal to linear trend. Comparison of T_2 with T_3 orthogonal to quadratic trend.

$$(b)\ \Sigma_1\xi_1' = 2, \quad \Sigma_1\xi_2' = 0,$$
$$\Sigma_2\xi_1' = -3, \quad \Sigma_2\xi_2' = 3,$$
$$\Sigma_3\xi_1' = 1, \quad \Sigma_3\xi_2' = -3.$$

Efficiency $\simeq 0{\cdot}9584$.

It is hoped to give a fuller list of the useful solutions in a later paper. The method of construction is straightforward. For example, the design given for three treatments, six replicates and cubic trend is found as follows. Since the design is to be symmetric, treatment differences are automatically orthogonal to ξ_1' and ξ_3' so that we need only consider ξ_2'. For the 'positive' positions we have the following values of ξ_2':

ξ_1'	1	3	5	7	9	11	13	15	17
ξ_2'	−40	−37	−31	−22	−10	5	23	44	68
(*a*)	T_1	T_2	T_2	T_3	T_1	T_3	T_3	T_1	T_2
(*b*)	T_1	T_2	T_1	T_3	T_2	T_3	T_3	T_2	T_1
(*c*)	T_1	T_2	T_3	T_1	T_2	T_3	T_3	T_2	T_1

	(*a*)	(*b*)	(*c*)
$\Sigma_1\xi_2'$	−12	−6	12,
$\Sigma_2\xi_2'$	0	−6	−6,
$\Sigma_3\xi_2'$	12	12	−6.

The problem is to select from these nine positions three for each treatment so that the sum of the values of ξ_2' over each treatment is as nearly zero as possible. A few minutes' trial shows that an exact solution is impossible and that the three designs above are the nearest solutions. The efficiency by (18) depends on the sum of squares of the $\Sigma_i\xi_2'$. Therefore (*b*) and (*c*) are the best solutions and by (18) their efficiency is 0·9992.

The designs given above for $p = 3$ are the best symmetric solutions. It is possible that more efficient solutions could be found by considering non-symmetric designs. The analysis of the non-symmetric designs would, however, be a little more complicated and the designs given above are, for most practical purposes, completely efficient.

A further useful property of some of the above designs is that certain treatment comparisons can be made with full accuracy even though the design as a whole is not fully efficient. For example, in design (*a*) for three treatments and three replicates, the comparison $(T_2 - T_3)$ is orthogonal to ξ_2 since $\Sigma_2\xi_2' = \Sigma_3\xi_2'$. Thus, for example, if T_1, T_2, T_3 are three levels of one treatment we can arrange that the linear part of 'between treatments' is fully accurate and orthogonal to the quadratic trend.

Some of the designs ($t = 3, n = 4$; $t = 4, n = 4$; $t = 3, n = 3$) turn out to be in blocks with one replicate of each treatment in each block, the arrangement of treatments within a block being systematic. Designs of this form were often advocated in the early literature of agricultural field trials. The present work shows that some of them are optimum provided that they are analysed correctly and that the assumptions on which the analysis is based are satisfied.

6. Comparison with other designs

The example in §4·2 showed that the selection of a design at random and the elimination of the trend by least squares (analysis of covariance) may lead to considerable loss of efficiency in estimating the treatment differences. It is interesting to find the average efficiency over all possible designs. Suppose, first, that only linear trends are involved ($p = 1$), so that the analysis of §4·3 holds. We have from (18) when $(t-2)\Sigma s_i^2/n(t-1)$ is small compared with one,

$$E \simeq 1 - \frac{\Sigma s_i^2}{n(t-1)}, \tag{22}$$

where $s_i = \Sigma_i \xi_1$. To average (22) over all designs we need the following lemma:

Lemma. *Let $N = tn$, where t, n are integers. Let $x_1, \ldots, x_N$ be N numbers such that $\Sigma x_i = 0$, $\Sigma x_i^2 = 1$. Divide the x_i at random into t blocks of n and let $s_1, \ldots, s_t$ be the sums of the x's in the separate blocks. Then the mean value of $s_1^2 + \ldots + s_t^2$ is $(N-n)/(N-1)$.*

The result follows on enumerating all arrangements, or can be deduced from formulae on sampling from finite populations.

Applying the Lemma to (22), we find that the average efficiency $\bar{E}$ is given by

$$\bar{E} \simeq 1 - \frac{1}{N-1} = \frac{N-2}{N-1}. \tag{23}$$

In the same way the approximate result (21) gives in the general case

$$\bar{E} \simeq \left(\frac{N-p-1}{N-1}\right). \tag{24}$$

Equations (23) and (24) are exactly true for two treatments. (These are average efficiencies: in unfavourable cases the efficiency may be appreciably less than this.)

In the example given in §4·2, $t = 2, N = 8, p = 3$ and $\bar{E} = 4/7 = 0{\cdot}5714$.

In general the gain in efficiency due to the use of systematic designs is only large when the order of polynomial is not too small compared with the number of plots.

Another instructive comparison is with randomized blocks, t plots per block. If there is a linear trend of amount $b\sigma$ per plot, then in the absence of treatment effects the measured quantity for the t plots in any block can be written

$$\text{const.} + \epsilon_1 + b\sigma, \text{const.} + \epsilon_2 + 2b\sigma, \ldots, \text{const.} + \epsilon_t + tb\sigma,$$

where $\epsilon_1, \ldots, \epsilon_t$ are independently distributed with mean zero and variance σ^2. Two plots are selected at random to carry treatments T_1 and T_2 and the average variance of the comparison of T_1 with T_2 is, for this block,

$$\mathscr{E}[\epsilon_i + ib\sigma - \epsilon_j - jb\sigma]^2, \tag{25}$$

where (i, j) form a random sample without replacement from $1, \ldots, t$. After some algebra (25) becomes

$$2\sigma^2\left[1 + \frac{t(t+1)\,b^2}{12}\right],$$

and for n blocks the variance is

$$\frac{2\sigma^2}{n}\left[1 + \frac{t(t+1)}{12}b^2\right],$$

giving an efficiency

$$E = \left(1 + \frac{t(t+1)}{12}b^2\right)^{-1}. \tag{26}$$

This shows, as would naturally be expected, that the steeper the trend and the larger the number of plots per block the less efficient is the randomized block design. These conclusions hold also for higher order trends.

To sum up, we have compared three ways of arranging the experiment:

(*a*) by a systematic design;

(*b*) by complete randomization and the use of covariance analysis to eliminate the trend; and

(*c*) by randomized blocks.

We can nearly always find a systematic design of efficiency very near unity, i.e. for which the variance of a treatment comparison is only slightly greater than $2\sigma^2/n$. The second method gives an average efficiency (24) independently of the amount of the trend. The third method gives an efficiency (26) decreasing as the amount of the trend increases.

For the analysis of the systematic designs to be valid we must be able to *assume* from our physical knowledge of the experiment that the measured quantity is formed by the addition of polynomial trend, treatment effect and completely random variation. This is not necessary when the design is randomized.

7. Analysis of results

In this section we show briefly how the results from one of these experiments should be analysed. The assumptions on which the analysis is based are stated at the end of §6. The method of analysis is a standard form of the method of least squares (analysis of covariance).

First, consider the exact designs for two treatments given in §3 of the paper. We can form an analysis of variance as follows. The treatment difference and sum of squares are obtained as if there was no trend. The trend coefficients and sums of squares are obtained as if there was no treatment difference. The residual sum of squares is obtained by subtraction:

	D.F.
Treatments	1
Trend	p
Residual	$2n-p-2$
Total	$2n-1$

The residual mean square s^2 estimates the error variance σ^2, and the estimated standard error of the treatment difference is $s(2/n)^{\frac{1}{2}}$. If the residual variation is normal the usual exact tests of significance hold.

In other cases the analysis is not as simple because trend and treatment effects are not exactly orthogonal. We can set the calculations out in the form of analysis of covariance or in the form of least squares theory. We can illustrate the latter method on the particular case of t treatments with cubic trend, the treatment differences being exactly orthogonal to ξ_1' and ξ_3'. We calculate sums of squares for ξ_1', ξ_2', and ξ_3' as if there were no treatment differences. We set up the least squares equations (13) for the joint estimation of the coefficient of ξ_2' and treatment effects. (The equations are in a form suitable for easy solution by iteration.) We

thus get both the best estimates of the treatment effects and the sum of squares for fitting ξ_2' and treatment constants. We have the analysis of variance:

	S.S.	D.F.
Fitting ξ_1' and ξ_3'	U_1	2
Fitting treatments and ξ_2'	U_3	t
Fitting ξ_2'	U_2	1
Treatments	U_3-U_2	$t-1$
Residual	$U_4-U_1-U_3$	$nt-t-3$
Total	U_4	$nt-1$

The residual mean square estimates the error variance σ^2, and the standard error of any treatment comparisons can be estimated directly from equation (16). If the distribution of the residual is normal, exact F and t tests can be made in the usual way.

8. Application of designs

It has been assumed that the order of trend p is given *a priori*. This is not normally true in practice, so that it is usually necessary to select a design with good properties for the largest p likely to occur. This value of p will depend on the nature of the experimental material and on the number of plots.

The solutions given above can be made the basis of more complicated designs. For example, the experiment may be repeated in a number of blocks, a trend being estimated from each block.

In the discussion it has been assumed that the main object of the experiment is to investigate the treatment differences. In some applications, such as that mentioned at the beginning of this paper, the trend itself is of direct physical interest. The same type of argument as used in §4 shows that the above systematic designs give optimum estimation of the trend.

Summary

Systematic experimental designs are given for use when a number of treatments are to be compared, one treatment being applied to each of a number of equally spaced plots. It is assumed that there is a smooth trend between plots and that the 'error' is independent in different plots. The designs enable the treatment effects (and the trend) to be estimated as simply and accurately as possible.

Sincere thanks are due to Mr R. C. Palmer, Mr F. J. Anscombe and Dr J. Wishart for helpful discussions.

REFERENCES

Cox, G. M. (1950). *Biometrics*, **6**, 317.

Fisher, R. A. & Yates, F. (1948). *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd.

Hald, A. (1948). *On the Decomposition of a Series of Observations.* Copenhagen: Gads Varlag.

Neyman, J. (1929). *The Theoretical Basis of Different Methods of Testing Cereals.* Part II. Warsaw: K. Buszczynski and Sons Ltd.

Williams, R. M. (1949). Unpublished Ph.D. Thesis, Cambridge University.

THE COMPARISON OF SEVERAL GROUPS OF OBSERVATIONS WHEN THE RATIOS OF THE POPULATION VARIANCES ARE UNKNOWN

By G. S. JAMES

University of Leeds

1. Introduction

B. L. Welch (1947) has described a method for comparing the means of two normal populations, when the ratio of their variances is unknown. The same method can be used to compare two regression coefficients. In the present paper we derive by similar methods a test for the equality of $k = (r+1)$ means, or regression coefficients.

Let $x_1, \ldots, x_k$ be any quantities which are distributed independently and normally with means $\mu_1, \ldots, \mu_k$ and variances $\alpha_1, \ldots, \alpha_k$ respectively. Let a_i $(i = 1, \ldots, k)$ be estimates of the α_i based on ν_i degrees of freedom, the a_i being distributed independently of the x_i and of each other in the forms

$$\frac{1}{\Gamma(\frac{1}{2}\nu_i)}\left(\frac{\nu_i a_i}{2\alpha_i}\right)^{\frac{1}{2}\nu_i - 1} \exp\left(-\frac{\nu_i a_i}{2\alpha_i}\right) d\left(\frac{\nu_i a_i}{2\alpha_i}\right). \tag{1}$$

(In the case of the comparison of means, the x_i will denote the means of the groups, while the α_i and a_i will denote the quantities called σ_i^2/n_i and s_i^2/n_i in the usual notation. If we are comparing the slopes of several regression lines, the x_i will denote the estimated slopes and the α_i the variances of these estimates. The principle can obviously be extended to more complicated situations, such as testing whether several regression lines have a common point for a specified value of the independent variate.)

If the α_i were known *a priori*, we could test the hypothesis that $\mu_1 = \ldots = \mu_k$ by using the fact that, when it is true,

$$\Sigma\omega_i(x_i - \bar{x})^2 \equiv \Sigma\omega_i x_i^2 - (\Sigma\omega_i x_i)^2/\omega \tag{2}$$

is distributed as χ^2 on $r = (k-1)$ degrees of freedom, where $\omega_i = 1/\alpha_i$, $\omega = \Sigma\omega_i$ and $\bar{x} = \Sigma\omega_i x_i/\omega$. That is to say,

$$\Pr\left[\Sigma\omega_i x_i^2 - (\Sigma\omega_i x_i)^2/\omega \leqslant 2\xi\right] = G_\rho(\xi), \tag{3}$$

where 2ξ denotes a particular value of χ^2 (such as the 5 % point), $\rho = \frac{1}{2}r = \frac{1}{2}(k-1)$ and

$$G_\rho(\xi) = \frac{1}{\Gamma(\rho)}\int_0^\xi t^{\rho-1} e^{-t} dt. \tag{4}$$

When the α_i are not known, but the ν_i are large, we may continue to use the test described in (2) or (3) as a 'large-sample' test, replacing ω_i by $w_i = 1/a_i$ and ω by $w = \Sigma w_i$. If the ν_i are not quite large, this procedure will not be very accurate, and a 'Studentized' test which takes into account the sampling variations of the a_i is desirable. Following Welch, we shall attempt to find a series development of a function $h(a_1, \ldots, a_k; \xi) \equiv h(a)$ of the variance estimates a_i and of ξ only, such that in place of (3) we shall have

$$\Pr\left[\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w \leqslant 2h(a)\right] = G_\rho(\xi). \tag{5}$$

It will be observed that on the left-hand side of the inequality in (5) we have replaced the true weights ω_i occurring in (2) and (3) by estimated weights w_i. The expression could therefore be written as $\Sigma w_i(x_i-\bar{x})^2$, where $\bar{x}$ now denotes $\Sigma w_i x_i/w$, but we shall not use this form, as it is likely to conceal the fact that $\bar{x}$ depends on the w_i.

2. PRELIMINARY ANALYSIS

From (5) we have $$\int \Pr\left[\Sigma w_i x_i^2-(\Sigma w_i x_i)^2/w \leqslant 2h(a) \mid a\right]\Pr\left[da\right] = G_\rho(\xi), \tag{6}$$

where the first expression denotes the probability of the relation indicated for *fixed* a_i (and hence for fixed w_i), and the second denotes the product of the probability differentials (1), the whole being integrated over the range 0 to ∞ of all the a_i. Now, by Taylor's theorem, we have the symbolic expansion

$$\Pr\left[\Sigma w_i x_i^2-(\Sigma w_i x_i)^2/w \leqslant 2h(a) \mid a\right] = \exp\left[\Sigma(a_i-\alpha_i)\,\partial_i\right]\Pr\left[\Sigma\omega_i x_i^2-(\Sigma\omega_i x_i)^2/\omega \leqslant 2h(\alpha)\right], \tag{7}$$

where, after expansion of the exponential, ∂_i denotes $\partial/\partial\alpha_i$. It should be carefully noted that the symbols ω_i and ω are used in this work merely as abbreviations for certain functions of the α_i, so that the operators ∂_i apply to them just as they do to the α_i appearing explicitly. It is also necessary to notice that the notation $\Pr[\ldots]$ used in (7) does not explicitly mention the distribution of the x_i, which itself depends on the α_i; this distribution is not, of course, supposed to be affected by the operators ∂_i, so that if we were using a fuller notation it might be necessary to use some device to distinguish between those α_i which were to be differentiated and those which were not. Substituting (7) into (6) we obtain

$$\begin{aligned} G_\rho(\xi) &= \int \exp\left[\Sigma(a_i-\alpha_i)\,\partial_i\right]\Pr\left[da\right]\Pr\left[\Sigma\omega_i x_i^2-(\Sigma\omega_i x_i)^2/\omega \leqslant 2h(\alpha)\right] \\ &= \Theta\Pr\left[\Sigma\omega_i x_i^2-(\Sigma\omega_i x_i)^2/\omega \leqslant 2h(\alpha)\right], \end{aligned} \tag{8}$$

where $$\begin{aligned} \Theta &= \Pi\left[e^{-\alpha_i\partial_i}\left(1-\frac{2\alpha_i\partial_i}{\nu_i}\right)^{-\frac{1}{2}\nu_i}\right] \\ &= 1+\Sigma\frac{\alpha_i^2\partial_i^2}{\nu_i}+\left\{\frac{4}{3}\Sigma\frac{\alpha_i^3\partial_i^3}{\nu_i^2}+\frac{1}{2}\left(\Sigma\frac{\alpha_i^2\partial_i^2}{\nu_i}\right)^2\right\}+O(\nu^{-3}). \end{aligned} \tag{9}$$

This result follows on substituting for $\Pr[da]$ from (1), exactly as in Welch's paper.

Now let $$\begin{aligned} h(a) &= h_0(a)+h_+(a) \\ &= h_0(a)+h_1(a)+h_2(a)+\ldots, \end{aligned} \tag{10}$$

where $h_s(a)$ is of order ν^{-s}. The $h_s(a)$ will be determined successively in the sequel. Introducing a further symbolic expansion, equation (8) now becomes

$$G_\rho(\xi) = \Theta\exp\left[h_+(\alpha)\,D\right]\Pr\left[\Sigma\omega_i x_i^2-(\Sigma\omega_i x_i)^2/\omega \leqslant 2h_0(\alpha)\right], \tag{11}$$

where, after expansion, D denotes differentiation with respect to the expression $h_0(\alpha)$ appearing on the right-hand side of the inequality—a function which will prove to be just ξ, as might be expected. Substituting for Θ and $h_+(\alpha)$ from (9) and (10) we obtain, to the second order,

$$G_\rho(\xi) = \left[1+\Sigma\frac{\alpha_i^2\partial_i^2}{\nu_i}+\left\{\frac{4}{3}\Sigma\frac{\alpha_i^3\partial_i^3}{\nu_i^2}+\frac{1}{2}\left(\Sigma\frac{\alpha_i^2\partial_i^2}{\nu_i}\right)^2\right\}\right]\left[1+h_1(\alpha)\,D+\left\{h_2(\alpha)\,D+\tfrac{1}{2}h_1^2(\alpha)\,D^2\right\}\right]\Pr\left[\ldots\right]. \tag{12}$$

Equating terms of successive orders, and using (3), we find

$$h_0(\alpha) = \xi, \tag{13}$$

$$\left[h_1(\alpha)\, D + \Sigma \frac{\alpha_i^2 \partial_i^2}{\nu_i}\right] \Pr\left[\Sigma \omega_i x_i^2 - (\Sigma \omega_i x_i)^2/\omega \leqslant 2\xi\right] = 0, \tag{14}$$

$$\left[h_2(\alpha)\, D + \tfrac{1}{2} h_1^2(\alpha)\, D^2 + \Sigma \frac{\alpha_i^2}{\nu_i} (h_1^{(i)}(\alpha) + 2h_1^{(i)}(\alpha)\, \partial_i + h_1(\alpha)\, \partial_i^2)\, D \right.$$
$$\left. + \frac{4}{3} \Sigma \frac{\alpha_i^3 \partial_i^3}{\nu_i^2} + \frac{1}{2} \left(\Sigma \frac{\alpha_i^2 \partial_i^2}{\nu_i}\right)^2 \right] \Pr[\ldots] = 0, \tag{15}$$

where $h_1^{(i)}(\alpha) = \partial_i^2 h_1(\alpha)$, etc.

It now only remains to carry out the differentiations indicated in (14) and (15) (remembering that $D = \partial/\partial\xi$) in order to obtain $h_1(\alpha), h_2(\alpha)$, and thence $h_1(a), h_2(a)$. At the corresponding point in the problem of comparing two means, the ∂_i operators only had to operate on simple functions, but here they have to operate on $\Pr[\Sigma \omega_i x_i^2 - (\Sigma \omega_i x_i)^2/\omega \leqslant 2\xi]$. This expression is a k-ple integral, and the operation may be thought of as differentiation, with respect to the boundary only, of the integral of the probability density function of the x_i throughout a region in k-space. The method adopted to evaluate these derivatives is to change the boundary slightly, expand the integral in powers of the quantities specifying this change, and obtain the derivatives by comparison with Taylor's expansion.

3. Evaluation of the derivatives

We have
$$\Pr[\Sigma \omega_i x_i^2 - (\Sigma \omega_i x_i)^2/\omega \leqslant 2\xi] = \Pr\left[\Sigma \frac{x_i^2}{\alpha_i} - \left(\Sigma \frac{x_i}{\alpha_i}\right)^2 \Big/ \Sigma \frac{1}{\alpha_i} \leqslant 2\xi\right]. \tag{16}$$

Now replace α_i by $\alpha_i(1+\epsilon_i)$ in (16); we then have

$$J \equiv \Pr\left[\Sigma \frac{x_i^2}{\alpha_i(1+\epsilon_i)} - \left(\Sigma \frac{x_i}{\alpha_i(1+\epsilon_i)}\right)^2 \Big/ \Sigma \frac{1}{\alpha_i(1+\epsilon_i)} \leqslant 2\xi\right]$$
$$= \left[1 + \Sigma \alpha_i \epsilon_i \partial_i + \frac{1}{2!} (\Sigma \alpha_i \epsilon_i \partial_i)^2 + \ldots\right] \Pr\left[\Sigma \frac{x_i^2}{\alpha_i} - \left(\Sigma \frac{x_i}{\alpha_i}\right)^2 \Big/ \Sigma \frac{1}{\alpha_i} \leqslant 2\xi\right], \tag{17}$$

by Taylor's theorem. But we also have

$$J = (2\pi)^{-\frac{1}{2}k} \Pi \alpha_i^{-\frac{1}{2}} \int_{R_0} \exp\left(-\tfrac{1}{2} \Sigma \frac{x_i^2}{\alpha_i}\right) dx, \tag{18}$$

where $\int \ldots dx$ is an abbreviation for $\int \ldots \int \ldots dx_1 \ldots dx_k$ and R_0 denotes the region

$$\Sigma \frac{x_i^2}{\alpha_i(1+\epsilon_i)} - \left(\Sigma \frac{x_i}{\alpha_i(1+\epsilon_i)}\right)^2 \Big/ \Sigma \frac{1}{\alpha_i(1+\epsilon_i)} \leqslant 2\xi. \tag{19}$$

Making the transformation $x_i = (2\alpha_i)^{\frac{1}{2}} y_i$, (18) becomes

$$J = \pi^{-\frac{1}{2}k} \int_R \exp(-\Sigma y_i^2)\, dy, \tag{20}$$

$$R \equiv \Sigma \frac{y_i^2}{1+\epsilon_i} - \left(\Sigma \frac{y_i}{\alpha_i^{\frac{1}{2}}(1+\epsilon_i)}\right)^2 \Big/ \Sigma \frac{1}{\alpha_i(1+\epsilon_i)} \leqslant \xi. \tag{21}$$

Tracing back the origin of the quadratic form in (21), or directly, it is seen to be of rank $r = k-1$. Let its matrix be $\mathbf{A}$, so that the inequality in (21) can be written as $\mathbf{y'Ay} \leqslant \xi$. We can now find an orthogonal transformation $\mathbf{y} = \mathbf{Lz}$ to new variables z_i, which is such that

$$\mathbf{L'AL} = \text{diag}\,(1+\eta_1, \ldots, 1+\eta_r, 0). \tag{22}$$

(20) now becomes

$$J = \pi^{-\frac{1}{2}k}\int_{\sum_1^k (1+\eta_i)z_i^2 \leqslant \xi} \exp\left(-\sum_1^k z_i^2\right) dz = \pi^{-\frac{1}{2}r}\int_{\sum_1^r (1+\eta_i)z_i^2 \leqslant \xi} \exp\left(-\sum_1^r z_i^2\right) dz, \tag{23}$$

on integrating out z_k; dz now denotes $dz_1 \dots dz_r$. Now make the further transformation $z_i = (1+\eta_i)^{-\frac{1}{2}} u_i$, and write

$$(1+\eta_i)^{-1} = 1-\zeta_i \quad (i = 1, \dots, r). \tag{24}$$

Then
$$J = \pi^{-\frac{1}{2}r}\Pi(1-\zeta_i)^{\frac{1}{2}}\int_{\Sigma u_i^2 \leqslant \xi} \exp\left[-\Sigma(1-\zeta_i)u_i^2\right] du$$

$$= \pi^{-\frac{1}{2}r}\Pi(1-\zeta_i)^{\frac{1}{2}} \sum_{b_1,\dots,b_r=0}^{\infty} \frac{\zeta_1^{b_1}\dots\zeta_r^{b_r}}{b_1!\dots b_r!}\int_{\Sigma u_i^2 \leqslant \xi} u_1^{2b_1}\dots u_r^{2b_r}\exp\left(-\Sigma u_i^2\right) du. \tag{25}$$

The integral is of Dirichlet's type and is easily evaluated by making a polar transformation, for example, yielding

$$J = \pi^{-\frac{1}{2}r}\Pi(1-\zeta_i)^{\frac{1}{2}}\,\Sigma\frac{\zeta_1^{b_1}\dots\zeta_r^{b_r}}{b_1!\dots b_r!}\frac{\Gamma(b_1+\frac{1}{2})\dots\Gamma(b_r+\frac{1}{2})}{\Gamma(\Sigma b_i+\frac{1}{2}r)}\int_0^{\xi} t^{\Sigma b_i+\frac{1}{2}r-1}e^{-t}\,dt$$

$$= \Pi(1-\zeta_i)^{\frac{1}{2}}\prod_{i=1}^{r}\frac{(\frac{1}{2})(\frac{3}{2})\dots(b_i-\frac{1}{2})}{b_i!}\zeta_i^{b_i}G_{\rho+\Sigma b_i}(\xi)$$

$$= \Pi\left(\frac{1-\zeta_i E}{1-\zeta_i}\right)^{-\frac{1}{2}} G_\rho(\xi), \tag{26}$$

where E is an operator with the property $E^b G_\rho(\xi) = G_{\rho+b}(\xi)$. We also define $\Delta = E - 1$. Then, using (24), (26) becomes

$$J = \prod_1^r (1-\eta_i\Delta)^{-\frac{1}{2}} G_\rho(\xi), \tag{27}$$

a remarkably simple symbolic form, but one which is unfortunately not so simple when expressed in terms of the ϵ_i.

Now, since $\mathbf{L}$ is orthogonal, (22) yields

$$\mathbf{L}'[(1+\Delta)\mathbf{I} - \Delta\mathbf{A}]\mathbf{L} = \operatorname{diag}(1-\eta_1\Delta, \dots, 1-\eta_r\Delta, 1+\Delta). \tag{28}$$

Taking determinants, and inserting the typical element of $\mathbf{A}$ by inspection of (21), we obtain

$$\begin{aligned}(1+\Delta)\Pi(1-\eta_i\Delta) &= |(1+\Delta)\mathbf{I} - \Delta\mathbf{A}| \\ &= \left|\left(1+\Delta-\frac{\Delta}{1+\epsilon_i}\right)\delta_{ij} + \frac{\Delta}{(\alpha_i\alpha_j)^{\frac{1}{2}}(1+\epsilon_i)(1+\epsilon_j)}\Big/\Sigma\frac{1}{\alpha_i(1+\epsilon_i)}\right| \\ &= |p_i\delta_{ij} + q_i q_j| \\ &= p_1\dots p_k\left(1+\sum_1^k q_i^2/p_i\right),\end{aligned} \tag{29}$$

where δ_{ij} is the Kronecker delta and

$$p_i = \frac{1+\epsilon_i(1+\Delta)}{1+\epsilon_i}, \quad q_i^2 = \frac{\Delta}{\alpha_i(1+\epsilon_i)^2}\Big/\Sigma\frac{1}{\alpha_i(1+\epsilon_i)}. \tag{30}$$

Substituting for p_i and q_i^2 in (29), we have

$$(1+\Delta)\prod_1^r(1-\eta_i\Delta) = \prod_1^k\left[\frac{1+\epsilon_i(1+\Delta)}{1+\epsilon_i}\right]\cdot\left[1+\Sigma\frac{\Delta}{\alpha_i(1+\epsilon_i)[1+\epsilon_i(1+\Delta)]}\Big/\Sigma\frac{1}{\alpha_i(1+\epsilon_i)}\right], \tag{31}$$

$$\prod_1^r(1-\eta_i\Delta)^{-\frac{1}{2}} = \prod_1^k\left[\frac{1+\epsilon_i(1+\Delta)}{1+\epsilon_i}\right]^{-\frac{1}{2}}\cdot\left[\frac{\Sigma\omega_i/[1+\epsilon_i(1+\Delta)]}{\Sigma\omega_i/(1+\epsilon_i)}\right]^{-\frac{1}{2}}. \tag{32}$$

It is now necessary to expand this expression in powers of the ϵ_i, and let it operate on $G_\rho(\xi)$ in accordance with (27). It is to be noted that

$$\Delta G_\rho(\xi) = -g_{\rho+1}(\xi), \tag{33}$$

where

$$g_\rho(\xi) = G'_\rho(\xi) = \frac{1}{\Gamma(\rho)}\xi^{\rho-1}e^{-\xi}. \tag{34}$$

Comparing the resulting expansion of J with that in (17), we finally obtain, after a good deal of algebra,

$$\alpha_i\partial_i \Pr[\ldots] = \frac{1}{2}\left(1-\frac{\omega_i}{\omega}\right)g_{\rho+1}(\xi), \tag{35}$$

$$\alpha_i^2\partial_i^2 \Pr[\ldots] = -\frac{1}{4}\left(1-\frac{\omega_i}{\omega}\right)^2(3E+1)g_{\rho+1}(\xi), \tag{36}$$

$$\alpha_i^3\partial_i^3 \Pr[\ldots] = \frac{3}{8}\left(1-\frac{\omega_i}{\omega}\right)^3(5E^2+2E+1)g_{\rho+1}(\xi), \tag{37}$$

$$\alpha_i^4\partial_i^4 \Pr[\ldots] = -\frac{3}{16}\left(1-\frac{\omega_i}{\omega}\right)^4(35E^3+15E^2+9E+5)g_{\rho+1}(\xi), \tag{38}$$

$$\begin{aligned}\alpha_i^2\alpha_j^2\partial_i^2\partial_j^2 \Pr[\ldots] = \Bigg[&-\frac{1}{16}\left(1-2\frac{\omega_i+\omega_j}{\omega}+\frac{\omega_i^2+\omega_j^2}{\omega^2}\right)(9E^3-3E^2-5E-1)\\ &-\frac{1}{4}\frac{\omega_i\omega_j}{\omega^2}(27E^3+3E^2+E+1)+\frac{1}{8}\frac{\omega_i^2\omega_j+\omega_i\omega_j^2}{\omega^3}(45E^3+9E^2+7E+3)\\ &-\frac{3}{16}\frac{\omega_i^2\omega_j^2}{\omega^4}(35E^3+15E^2+9E+5)\Bigg]g_{\rho+1}(\xi)\quad(i\neq j).\end{aligned} \tag{39}$$

It may be noticed that (39) does not reduce to (38) when $i=j$.

4. Completion of solution

Substituting (36) into (14) we obtain

$$h_1(\alpha) = \frac{1}{4}\left(3\frac{\xi^2}{\rho(\rho+1)}+\frac{\xi}{\rho}\right)\Sigma\frac{1}{\nu_i}\left(1-\frac{\omega_i}{\omega}\right)^2. \tag{40}$$

Differentiating (40) twice with respect to α_i, and substituting these results, together with (35)–(39), into (15), we finally obtain, as the approximation of order -2 in the ν_i, the expression

$$\begin{aligned}2h(a) = \chi^2 &+ \tfrac{1}{2}(3\chi_4+\chi_2)\Sigma\frac{1}{\nu_i}\left(1-\frac{w_i}{w}\right)^2+\left\{\tfrac{1}{16}(3\chi_4+\chi_2)^2\left(1-\frac{k-3}{\chi^2}\right)\left(\Sigma\frac{1}{\nu_i}\left(1-\frac{w_i}{w}\right)^2\right)^2\right.\\ &+\tfrac{1}{2}(3\chi_4+\chi_2)[(8R_{23}-10R_{22}+4R_{21}-6R_{12}^2+8R_{12}R_{11}-4R_{11}^2)\\ &+(2R_{23}-4R_{22}+2R_{21}-2R_{12}^2+4R_{12}R_{11}-2R_{11}^2)(\chi_2-1)\\ &+\tfrac{1}{4}(-R_{12}^2+4R_{12}R_{11}-2R_{12}R_{10}-4R_{11}^2+4R_{11}R_{10}-R_{10}^2)(3\chi_4-2\chi_2-1)]\\ &+(R_{23}-3R_{22}+3R_{21}-R_{20})(5\chi_6+2\chi_4+\chi_2)\\ &+\tfrac{3}{16}(R_{12}^2-4R_{23}+6R_{22}-4R_{21}+R_{20})(35\chi_8+15\chi_6+9\chi_4+5\chi_2)\\ &+\tfrac{1}{16}(-2R_{22}+4R_{21}-R_{20}+2R_{12}R_{10}-4R_{11}R_{10}+R_{10}^2)(9\chi_8-3\chi_6-5\chi_4-\chi_2)\\ &+\tfrac{1}{4}(-R_{22}+R_{11}^2)(27\chi_8+3\chi_6+\chi_4+\chi_2)\\ &\left.+\tfrac{1}{4}(R_{23}-R_{12}R_{11})(45\chi_8+9\chi_6+7\chi_4+3\chi_2)\right\}+O(\nu^{-3}).\end{aligned} \tag{41}$$

In this expression we write $$R_{st} = \Sigma \frac{1}{\nu_i^s} \left(\frac{w_i}{w}\right)^t, \tag{42}$$

while in order to facilitate printing χ_{2s} is written in place of $\chi^{2s}/(k-1)(k+1)\dots(k+2s-3)$.

If we only wish to go as far as the term of order -1 in the ν_i, this result may conveniently be written

$$2h(a) = \chi^2\left[1 + \frac{3\chi^2+(k+1)}{2(k^2-1)} \Sigma \frac{1}{\nu_i}\left(1-\frac{w_i}{w}\right)^2 + O(\nu^{-2})\right]. \tag{43}$$

To summarize, then, to test the significance of the differences between $x_1, \dots, x_k$, we calculate the statistic

$$\Sigma w_i x_i^2 - (\Sigma w_i x_i)^2/w = \chi_c^2, \text{ say.} \tag{44}$$

If the ν_i are all large, this can be taken to have the χ^2 distribution with $(k-1)$ degrees of freedom, so that if χ^2 denotes the value appropriate to a given probability P, then $\Pr[\chi_c^2 \leqslant \chi^2] = P$. However, if the ν_i are not all large, we must not compare χ_c^2 with the percentage point χ^2, but with another *statistic* $2h(a)$, of which (41) is an expansion in inverse powers of the ν_i.

The actual manner in which (41) 'converges to the true function $2h(a)$' is not known, but it is hoped that the use of at any rate the first corrective term (incorporated in (43)) may result in a test which is a good deal more accurate than the large sample χ^2 test, at any rate for moderately large ν_i. This is known to be so in the case $k = 2$ considered by Welch. (See, for example, Aspin (1948), table 1.) In any case, to include the second corrective term would involve a good deal of numerical calculation.

It should, perhaps, be mentioned that the result (41) can be shown to agree with Welch's result when $k = 2$. In this demonstration it has to be noted that certain non-linear relations hold between the quantities R_{st} in the special case $k = 2$.

The author wishes to express his thanks to Mr A. M. Walker for much helpful discussion on different methods of dealing with problems of the kind considered in this paper. It is hoped to publish in a subsequent paper methods for dealing with more general linear hypotheses than the one considered here, and also with some multivariate problems.

REFERENCES

ASPIN, ALICE A. (1948). An examination and further development of a formula arising in the problem of comparing two mean values. *Biometrika*, **35**, 88–96.

WELCH, B. L. (1947). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, **34**, 28–35.

Note added in proof. It should be pointed out that (5) is by no means the only possible way of modifying (3) to obtain a test for smaller samples. We could attempt to find functions

$$W_i(a) = 1/[a_i + j_{i1}(a) + \dots]$$

such that

$$\Pr[\Sigma W_i x_i^2 - (\Sigma W_i x_i)^2/W \leqslant 2\xi] = G_\rho(\xi),$$

the j_{i1} being $O(\nu^{-1})$. It then transpires that the j_{i1} have to satisfy the single equation

$$\Sigma\left[\left(1-\frac{\omega_i}{\omega}\right)\frac{j_{i1}(\alpha)}{\alpha_i} - \frac{3\chi^2+(k+1)}{2(k+1)}\frac{1}{\nu_i}\left(1-\frac{\omega_i}{\omega}\right)^2\right] = 0.$$

If we look for a solution of the form $j_{i1}(\alpha) = \alpha_i l(\alpha)$, we are led to our original test (43). Another obvious solution is obtained by setting each term under the summation sign equal to zero.

ON THE COMPARISON OF SEVERAL MEAN VALUES: AN ALTERNATIVE APPROACH

By B. L. WELCH
University of Leeds

1. Introduction

Let y_t $(t = 1, 2, \ldots, k)$ be statistical quantities normally and independently distributed with means μ_t and variances $\lambda_t \sigma_t^2$, respectively, where the λ_t are known constants but nothing is known about μ_t and σ_t^2. Suppose that the data provide estimates s_t^2 of the σ_t^2 which are distributed, respectively, as $\chi_t^2 \sigma_t^2 / f_t$ where f_t is the number of degrees of freedom of χ_t^2. Suppose, furthermore, that all the s_t^2 are distributed independently of each other and of all the y_t. The problem considered here is then to test whether the data are consistent with the hypothesis that all the μ_t are equal, i.e. $\mu_t = \mu$ $(t = 1, 2, \ldots, k)$.

As a particular case y_t may be the means $\bar{x}_t$ of samples of n_t observations from k different normal populations whose true means and variances are μ_t and σ_t^2, respectively. Since $V(\bar{x}_t)$ is σ_t^2 / n_t we have $\lambda_t = 1/n_t$. The separate samples provide estimates s_t^2 of the σ_t^2. The hypothesis under review is then whether the populations sampled may be considered to have the same mean without, however, imposing the condition that their variances are necessarily equal.

The problem has already been discussed by G. S. James (1951) elsewhere in the present journal (pp. 324–9). In my notation, which is slightly different from his, he considers the statistic $\Sigma w_t (y_t - \hat{y})^2$, where $w_t = 1/\lambda_t s_t^2$ and $\hat{y} = (\Sigma w_t y_t)/(\Sigma w_t)$. When all the μ_t are equal this quantity is, in large samples (i.e. provided f_t are large), distributed as χ^2 with $(k-1)$ degrees of freedom. If χ_P^2 is the value of χ^2 exceeded with probability P, we shall then have

$$\Pr\left[\sum_t w_t (y_t - \hat{y})^2 > \chi_P^2\right] = P. \tag{1}$$

For samples not large enough for this to apply James considers the problem of finding a function $h(w_1, w_2, \ldots, w_k, P)$, which may be written $h(w)$ for brevity, and which has the property that

$$\Pr\left[\sum_t w_t (y_t - \hat{y})^2 > h(w)\right] = P. \tag{2}$$

Hé develops a method of deriving successive approximations to such a function bringing in, in turn, terms of order $1/f_t$, $1/f_t^2$ and so on. To order $1/f_t$, for instance, he finds

$$h(w) = \chi_P^2 \left[1 + \frac{3\chi_P^2 + (k+1)}{2(k^2-1)} \sum_t \frac{1}{f_t} \left(1 - \frac{w_t}{\Sigma w_t}\right)^2\right]. \tag{3}$$

James also gives explicitly the term of order $1/f_t^2$ but, as he points out, this is so complicated as to discourage the view that it is likely to prove of much practical utility; nor does it appear to be worth while to carry through the heavy algebra needed to produce the terms of still higher orders. However, a procedure based only on the terms to order $1/f_t$ is quite easy to apply and, provided the f_t are not too small, is likely to be an appreciable improvement on the large-sample assumption that $\sum_t w_t (y_t - \hat{y})^2$ follows a χ^2 distribution. It may therefore be of interest to derive the result given in (3) by an alternative route. This I propose

to do in the present paper. More specifically, I shall first develop an approximation which involves the use of the variance ratio tables rather than the χ^2 tables. But it will be seen that, to the order $1/f_i$, this is equivalent to the procedure given by James.

2. THE SPECIAL CASE, $k = 2$

Before proceeding it will be convenient to recall certain results for the particular case $k = 2$ which has already been discussed fairly fully by myself (Welch, 1947, 1949) and Aspin (1948, 1949). In this case we have

$$\sum_i w_i(y_i-\hat{y})^2 = \frac{w_1 w_2(y_1-y_2)^2}{(w_1+w_2)} = \frac{(y_1-y_2)^2}{(\lambda_1 s_1^2+\lambda_2 s_2^2)}, \tag{4}$$

i.e. $$\sum_i w_i(y_i-\hat{y})^2 = v^2, \quad \text{where} \quad v = (y_1-y_2)/\sqrt{(\lambda_1 s_1^2+\lambda_2 s_2^2)}.$$

Now when $\mu_1 = \mu_2$ the numerator of v is a normal deviate with standard deviation $\sqrt{(\lambda_1\sigma_1^2+\lambda_2\sigma_2^2)}$. The denominator of v is an estimate of this standard deviation obtained by replacing σ_1^2 and σ_2^2 by the sample variances s_1^2 and s_2^2. Welch and Aspin have made some progress towards tabling critical values of a quantity v defined in this way. At the same time, however, I have also given some consideration to an alternative approach which requires only the already available tables of the 'Student' distribution, and which furnishes an approximation which numerical work suggests to be very useful. This alternative approach stems from the replacement of the true distribution of $(\lambda_1 s_1^2+\lambda_2 s_2^2)$ by a Pearson Type III curve with moments suitably adjusted. I was able to show (Welch, 1947) that the errors involved in making probability statements about v on the basis of this particular approximation were of order $1/f_i^2$.

The use of standard types of frequency curves to approximate to awkward sampling distributions is, of course, a common device. It is natural, therefore, to ask whether it has any useful application in the more general problem of making probability statements about $\sum_i w_i(y_i-\hat{y})^2$ when k may exceed 2. It is to this question that we now turn.

3. THE MOMENT-GENERATING FUNCTION OF $\sum_i w_i(y_i-\hat{y})^2$

We have seen in equation (4), above, that, when $k = 2$, $\sum_i w_i(y_i-\hat{y})^2$ separates into a pure function of the y_i divided by a pure function of the s_i^2. When $k > 2$ a like separation does not take place. It appears, therefore, that we have to approximate the distribution of $\sum_i w_i(y_i-\hat{y})^2$ directly without being able to make it depend on a function of the s_i^2, alone, which has simple moments. Now the moments of $\sum_i w_i(y_i-\hat{y})^2$ are not simple and, indeed, after a certain order they become infinite. Nevertheless, we shall proceed formally as if the moment-generating function, $M(u)$, of $\sum_i w_i(y_i-\hat{y})^2$ existed, writing

$$M(u) = E\exp[u\sum_i w_i(y_i-\hat{y})^2], \tag{5}$$

where E denotes averaging over the joint distributions of y_i and s_i^2. This can also be written

$$M(u) = E_1 E_2 \exp[u\sum_i w_i(y_i-\hat{y})^2], \tag{6}$$

where E_2 denotes averaging over the joint distribution of the s_i^2 and E_1 denotes averaging over the joint distribution of the y_i.

It will be recalled that $w_i = 1/\lambda_i s_i^2$. Similarly, we can denote $1/\lambda_i \sigma_i^2$ by ω_i. With this notation, expanding symbolically, we have

$$E_2 \exp[u \sum_i w_i (y_i - \hat{y})^2] = \Phi \exp[u \sum_i z_i (y_i - \bar{y})^2], \tag{7}$$

where

$$\Phi = E_2 \exp[\sum_i (w_i - \omega_i) D_i], \tag{8}$$

$$D_i^r = \left(\frac{\partial^r}{\partial z_i^r}\right)_{z=\omega} \quad \text{and} \quad \bar{y} = (\Sigma z_i y_i)/(\Sigma z_i).$$

Before Φ can be interpreted it must, of course, be expanded in powers of D_i. This gives an infinite series in D_i, but if we proceed only to a certain order of $1/f_i$, we have a series which terminates. Thus to the first order it is seen that

$$\Phi = 1 + 2\sum_i \frac{\omega_i D_i}{f_i} + \sum_i \frac{\omega_i^2 D_i^2}{f_i}. \tag{9}$$

In the present paper we shall work only to this order. Carrying through the differentiations we then obtain

$$\Phi \exp[u \sum_i z_i (y_i - \bar{y})^2] = \exp[u \sum_i \omega_i (y_i - \bar{y})^2]$$
$$\times \left[1 + 2u \sum_i \frac{\omega_i (y_i - \bar{y})^2}{f_i} - \frac{2u}{(\Sigma\omega_i)} \sum_i \omega_i^2 (y_i - \bar{y})^2 / f_i + u^2 \sum_i \omega_i^2 (y_i - \bar{y})^4 / f_i \ldots\right], \tag{10}$$

where, now, $\bar{y} = (\sum_i \omega_i y_i)/(\sum_i \omega_i)$. We then see from equations (6) and (7) that, in order to obtain $M(u)$, we must average the right-hand side of (10) over the joint distribution of the y_i.

Since we are at the moment only concerned with what happens when the μ_i are equal, and since (10) does not depend on the origin of co-ordinates, we lose no generality if we take $\mu_i = 0$, i.e.

$$p(y_1, y_2, \ldots, y_k) = (2\pi)^{-\frac{1}{2}k} (\omega_1 \omega_2 \ldots \omega_k)^{\frac{1}{2}} e^{-\frac{1}{2}\Sigma\omega_i y_i^2}. \tag{11}$$

We then have from (10) and (11)

$$M(u) = \int (2\pi)^{-\frac{1}{2}k} (\omega_1 \omega_2 \ldots \omega_k)^{\frac{1}{2}} \exp[-\tfrac{1}{2} \sum_i \omega_i y_i^2 + u \sum_i \omega_i (y_i - \bar{y})^2]$$
$$\times \left[1 + 2u \sum_i \frac{\omega_i}{f_i}\left(1 - \frac{\omega_i}{\Sigma\omega_i}\right)(y_i - \bar{y})^2 + u^2 \sum_i \omega_i^2 (y_i - \bar{y})^4 / f_i\right] dy_1 \ldots dy_k. \tag{12}$$

Using again the symbolic expression for Taylor's series this yields

$$M(u) = \Psi\left[1 + 2u \sum_i \frac{\omega_i}{f_i}\left(1 - \frac{\omega_i}{\Sigma\omega_i}\right) v_i^2 + u^2 \sum_i \frac{\omega_i^2 v_i^4}{f_i}\right], \tag{13}$$

where

$$\Psi = \int (2\pi)^{-\frac{1}{2}k} (\omega_1 \omega_2 \ldots \omega_k)^{\frac{1}{2}} \exp[-\tfrac{1}{2} \sum_i \omega_i y_i^2 + u \sum_i \omega_i (y_i - \bar{y})^2 + \sum_i (y_i - \bar{y})\, \delta_i] dy_1 \ldots dy_k, \tag{14}$$

and δ_i denotes differentiation with respect to v with v put equal to zero when all differentiations have been carried out. On integration

$$\Psi = (1 - 2u)^{-\frac{1}{2}(k-1)} \exp\left[\tfrac{1}{2}(1 - 2u)^{-1}\left\{\sum_i \frac{\delta_i^2}{\omega_i} - \frac{(\Sigma\delta_i)^2}{(\Sigma\omega_i)}\right\}\right], \tag{15}$$

it being understood that before this can be interpreted the exponent must be expanded in powers of δ_t. In virtue of the nature of the expression on which Ψ is to operate only a few of the terms in this expansion are needed. We find that

$$\Psi v_t^2 = (1-2u)^{-\frac{1}{2}(k-1)}\left[(1-2u)^{-1}\left\{\frac{1}{\omega_t}-\frac{1}{(\Sigma\omega_t)}\right\}\right], \tag{16}$$

$$\Psi v_t^4 = (1-2u)^{-\frac{1}{2}(k-1)}\left[3(1-2u)^{-2}\left\{\frac{1}{\omega_t}-\frac{1}{(\Sigma\omega_t)}\right\}^2\right], \tag{17}$$

whence from (13)

$$M(u) = (1-2u)^{-\frac{1}{2}(k-1)}\left[1+\{2u(1-2u)^{-1}+3u^2(1-2u)^{-2}\}\left\{\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2\right\}\right]. \tag{18}$$

Working, therefore, to order $1/f_t$, the corresponding cumulant-generating function of $\sum_t w_t(y_t-\bar{y})^2$ is

$$K(u) = -\tfrac{1}{2}(k-1)\log(1-2u)+\{2u(1-2u)^{-1}+3u^2(1-2u)^{-2}\}\left\{\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2\right\}. \tag{19}$$

4. Graduation by a Pearson type curve

At this stage we may, if we wish, invert (18) term by term to produce a distribution function, but at the present moment we are more concerned to see if one of the Pearson curves can be used in this connexion. The obvious one to try is that followed by the variance ratio F, since the F distribution is the one used in the simpler problem of comparing mean values when the ratios of the population σ_t's are known. Let us, therefore, consider the moment-generating function of the F distribution. (Again we shall proceed formally, since strictly the moments of F after a certain order will become infinite.)

We have $F = (\chi_1^2/\hat{f}_1) \div (\chi_2^2/\hat{f}_2)$, where χ_1^2 and χ_2^2 are distributed independently as χ^2 with degrees of freedom $\hat{f}_1$ and $\hat{f}_2$ respectively. Now

$$E\, e^{u\chi_1^2/\hat{f}_1} = \{1-2u/\hat{f}_1\}^{-\frac{1}{2}\hat{f}_1}. \tag{20}$$

Therefore, for *given* χ_2^2, we have $\quad E\, e^{uF} = \left\{1-\dfrac{2u\hat{f}_2}{\hat{f}_1\chi_2^2}\right\}^{-\frac{1}{2}\hat{f}_1}$. (21)

The moment-generating function, $M_F(u)$, is then given by averaging (21) over the χ_2^2 distribution. This does not give anything simple, but at this point it should be noted that, if the F distribution is going to be useful, we would expect from large-sample considerations that $\hat{f}_2$ would have to be of the same order of magnitude as the f_t of the earlier sections, and that $\hat{f}_1$ would differ from $(k-1)$ at most by a quantity of the order $1/\hat{f}_2$. Let us write down $M_F(u)$, therefore, correct only to order $1/\hat{f}_2$. Doing this we have, after averaging (21),

$$M_F(u) = (1-2u/\hat{f}_1)^{-\frac{1}{2}\hat{f}}\left[1+\frac{2u}{\hat{f}_2}(1-2u/\hat{f}_1)^{-1}+\frac{(\hat{f}_1+2)}{\hat{f}_1\hat{f}_2}u^2(1-2u/\hat{f}_1)^{-2}\right]. \tag{22}$$

Next let us write $\hat{f}_1 = (k-1)^*$ and let $G = [(k-1)+A/\hat{f}_2]F$. Then to order $1/\hat{f}_2$ we obtain, from (22),

$$M_G(u) = (1-2u)^{-\frac{1}{2}(k-1)}\left[1+\frac{(A+2\sqrt{k-1})}{\hat{f}_2}u(1-2u)^{-1}+\frac{(k^2-1)}{\hat{f}_2}u^2(1-2u)^{-2}\right], \tag{23}$$

i.e. $\quad K_G(u) = -\tfrac{1}{2}(k-1)\log(1-2u)+\dfrac{1}{\hat{f}_2}(A+2\sqrt{k-1})u(1-2u)^{-1}+\dfrac{(k^2-1)}{\hat{f}_2}u^2(1-2u)^{-2}$. (24)

* $f_1 = (k-1)+B/f_2$ was also tried but does not lead to a more general result. It is seen that B must equal zero.

Comparing this with the cumulant-generating function of $\sum_t w_t(y_t-\hat{y}_t)^2$ given in the last section, we see that we can bring the two into agreement if we write

$$\left.\begin{aligned}\frac{(k^2-1)}{f_2'} &= 3\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2 \\ \text{and}\quad \frac{(A+2\overline{k-1})}{f_2'} &= 2\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2,\end{aligned}\right\} \tag{25}$$

i.e.

$$\left.\begin{aligned}\frac{1}{f_2'} &= \frac{3}{(k^2-1)}\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2 \\ \text{and}\quad \frac{A}{f_2'} &= \frac{2(k-2)}{(k+1)}\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2.\end{aligned}\right\} \tag{26}$$

To order $1/f_t$ it appears, therefore, that $\Sigma w_t(y_t-\hat{y})^2$ is distributed as $\{(k-1)+A/f_2'\}$ times F, where $f_1' = (k-1)$, and f_2' and A/f_2' are given by (26).

5. Application to problem of testing $\mu_t = \mu$

The result in the last section may be restated as follows. Suppose we define a criterion

$$v^2 = \frac{\sum_t w_t(y_t-\hat{y})^2/(k-1)}{\left[1+\frac{2(k-2)}{(k^2-1)}\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2\right]}, \tag{27}$$

and suppose F_P is the tabled value of the variance ratio F, exceeded with probability P, for degrees of freedom $f_1' = (k-1)$ and

$$f_2' = \left[\frac{3}{(k^2-1)}\sum_t \frac{1}{f_t}\left(1-\frac{\omega_t}{\Sigma\omega_t}\right)^2\right]^{-1}.$$

Then, when $\mu_t = \mu$, we have approximately (i.e. to order $1/f_t$)

$$\Pr(v^2 > F_P) = P. \tag{28}$$

It will be seen from (27) that v^2 involves the unknown quantities $\omega_t\,(=1/\lambda_t\sigma_t^2)$ as well as the sample values $w_t\,(=1/\lambda_t s_t^2)$. The unknown ω_t also enter into f_2'. We cannot, therefore, use (28), as it stands, to furnish a test which can be applied in practice. However, since the ω_t, in each case, enter only into expressions of order $1/f_t$, the substitution of w_t for ω_t where they occur will still leave us with a probability statement (28) which is true to this order. Our approximate test procedure will, therefore, be:

(i) *Calculate*

$$v^2 = \frac{\sum_t w_t(y_t-\hat{y})^2/(k-1)}{\left[1+\frac{2(k-2)}{(k^2-1)}\sum_t \frac{1}{f_t}\left(1-\frac{w_t}{\Sigma w_t}\right)^2\right]}, \tag{29}$$

$$f_1' = (k-1);\quad f_2' = \left[\frac{3}{(k^2-1)}\sum_t \frac{1}{f_t}\left(1-\frac{w_t}{\Sigma w_t}\right)^2\right]^{-1}. \tag{30}$$

(ii) *Refer v^2 to a variance ratio table entered with degrees of freedom f_1' and f_2'.*

The procedure can, of course, be put into a form in which the Incomplete Beta Function Tables are used if it is felt that reference to a few percentage points is not adequate.

6. EXAMPLE

The data in Table 1 relate to an experiment in which three treatments are being compared, a single characteristic x being measured. As mentioned in §1 above, the y_t are now the means $\bar{x}_t$ for the three treatments. The s_t^2 are the individual 'within-treatment' variances, estimated on degrees of freedom f_t, equal to one less than the number of replicates in each case. Also $\lambda_t = 1/n_t$. Hence $w_t = n_t/s_t^2$.

Table 1

Treatment (t)	No. of individuals (n_t)	Treatment mean ($\bar{x}_t$)	Observed variance (s_t^2)	Estimated variance of mean (s_t^2/n_t)	w_t (n_t/s_t^2)
1	20	27·8	60·1	3·00	0·333
2	10	24·1	6·3	0·63	1·587
3	10	22·2	15·4	1·54	0·649
					2·569

If the true means μ_t are all equal, the best estimate of the common μ will be given by $\hat{x} = (\Sigma w_t \bar{x}_t)/(\Sigma w_t) = 24{\cdot}10$. The details required in this calculation and further quantities needed later are shown in Table 2. An arbitrary origin, $x = 25$, is being used.

Table 2

t	$w_t(\bar{x}_t - 25)$	$w_t(\bar{x}_t - 25)^2$	$w_t/\Sigma w_t$	$(1 - w_t/\Sigma w_t)^2$	$1/f_t(1 - w_t/\Sigma w_t)^2$
1	0·932	2·612	0·130	0·757	0·0398
2	−1·428	1·285	0·618	0·146	0·0162
3	−1·817	5·088	0·253	0·558	0·0620
	−2·313	8·985			0·1180

We have

$$\sum_t w_t(\bar{x}_t - \hat{x})^2 = 8{\cdot}985 - (2{\cdot}313)^2/2{\cdot}569 = 6{\cdot}90. \tag{31}$$

Substitution into (29) then gives

$$v^2 = \frac{\frac{1}{2}(6{\cdot}90)}{[1 + \frac{1}{4}(0{\cdot}1180)]} = \frac{3{\cdot}45}{1{\cdot}029} = 3{\cdot}35. \tag{32}$$

This must be referred to the variance-ratio table entered with $f_1 = 2$, and

$$f_2 = [\tfrac{3}{8}(0{\cdot}1180)]^{-1} = 8/(0{\cdot}354) = 22{\cdot}6. \tag{33}$$

The 5 % point for F, corresponding to these numbers of degrees of freedom, is 3·43. If, therefore, this point is taken as critical, the experiment just fails to demonstrate that the true means μ_t differ.

7. Equivalence with James's result to order $1/f_t$

The procedure described in equations (29) and (30) involves the use of F tables; the procedure due to James, described in equations (2) and (3) of §1, involves only the use of the χ^2 tables. It remains to show that the two procedures are equivalent to order $1/f_t$. To do this an expression is first required for the percentage points of F in terms of those of χ^2. Such an expression can be found in the same way as one finds an expansion for a t-deviate in terms of a normal deviate. If F has degrees of freedom f_1 and f_2, then, to order $1/f_2$, it may be shown that

$$F_P = \frac{\chi^2_P}{f_1}\left[1+\frac{\chi^2_P-(f_1-2)}{2f_2}\right], \tag{34}$$

where F_P is the value of F exceeded with probability P and χ^2_P is the corresponding value of χ^2 with f_1 degrees of freedom.

Using (34) in conjunction with (29) and (30) my procedure is seen, to order $1/f_t$, to be equivalent to taking as the critical value of v^2

$$v^2_P = \frac{\chi^2_P}{(k-1)}\left[1+\frac{3\{\chi^2_P-(k-3)\}}{2(k^2-1)}\sum_t\frac{1}{f_t}\left(1-\frac{w_t}{\Sigma w_t}\right)^2\right], \tag{35}$$

or, in other words, to taking as the critical value of $\sum_t w_t(y_t-\hat{y})^2$, the expression

$$h(w) = \chi^2_P\left[1+\frac{3\chi^2_P+(k+1)}{2(k^2-1)}\sum_t\frac{1}{f_t}\left(1-\frac{w_t}{\Sigma w_t}\right)^2\right]. \tag{36}$$

But this is just what is given in equation (3) of §1. Hence my procedure is equivalent to James's to order $1/f_t$.

To higher orders, of course, the two procedures differ. There are obvious points which could be mentioned in favour of either method, but no extensive numerical work has been done to compare them. We may note that, applied to the example in §6 above, equation (36) would give for the 5 % point of $\sum_t w_t(\bar{x}_t-\hat{x})^2$,

$$h(w) = 5{\cdot}991[1+\tfrac{1}{18}\{3(5{\cdot}991)+4\}(0{\cdot}118)] = 6{\cdot}96. \tag{37}$$

The observed value of $\sum_t w_t(\bar{x}_t-\hat{x})^2$ is 6·90, which is again just below the 5 % point, so that practically the same conclusion would be drawn.

REFERENCES

Aspin, A. A. (1948). *Biometrika*, **35**, 88–96.
Aspin, A. A. (1949). *Biometrika*, **36**, 290–3.
James, G. S. (1951). *Biometrika*, **38**, 324–9.
Welch, B. L. (1947). *Biometrika*, **34**, 28–35.
Welch, B. L. (1949). *Biometrika*, **36**, 293–6.

THE EFFECT OF STANDARDIZATION ON A χ^2 APPROXIMATION IN FACTOR ANALYSIS

By M. S. BARTLETT
University of Manchester

(With an Appendix by W. Ledermann)

1. *Explanation of the problem.* In a previous paper (Bartlett, 1950) I proposed a χ^2 analysis to test the significance of the successive roots of the determinantal equation

$$|R-\lambda| = 0, \tag{1}$$

where R is a correlation matrix of p variables; equation (1) is familiar in the particular method of factor analysis known as an analysis into principal components. An independent derivation of the χ^2 approximation for the significance of the product of *all* the roots,

$$\lambda_1\lambda_2 \dots \lambda_p = |R|, \tag{2}$$

was given by Box (1949) in this journal, in a systematic study of χ^2 and variance ratio approximations to certain likelihood criteria. In my own contribution, however, I was also concerned with the successive reduction of the total χ^2, corresponding to the elimination of successive components, analogously to the reduction of the total χ^2 in a canonical regression analysis between two groups of variables (Bartlett, 1938).

The validity of the χ^2 (or alternative) approximation for such reductions is more difficult to investigate in detail, as I have previously pointed out, for it depends on the non-null sampling distribution. Fortunately, it has still been possible from geometrical considerations to justify the approximation under appropriate conditions. Thus in the canonical analysis the elimination of *real* canonical components tends to reduce automatically the dimensions (degrees of freedom) of the observations (n-space), the dependent variables (p-space) and the independent variables (q-space). The appropriate conditions are that the components eliminated should be sufficiently well-determined in direction; this will be achieved even for small real effects as n increases, but alternatively, for moderate-sized n, as the real effects increase.

In the principal components case there are at least two variants of the problem of which the commoner is unfortunately the more difficult. In the simpler case equation (1) is replaced by the unstandardized variance-covariance matrix, the true variances being assumed equal. If the true common variance were also given, the problem would be a special case of the 'external' regression problem; but in the purely 'internal' tests for significance of the ratios of the roots, we may conveniently divide all the roots by their mean, which is identical with the mean observed variance. The geometrical argument may be immediately applied to this case, but the further effect of standardization, while readily allowed for in the total χ^2, is much less clear-cut when some components are eliminated, as I was obliged to note in my previous paper. Following a more detailed examination of the case of one 'real' (i.e. non-homogeneous) principal component, I have since added an amending note (Bartlett, 1951) summarizing the further conclusions reached. It seems desirable, however, to put on record also the details of this investigation, so that the precise basis of these further conclusions is available; this is done in the present paper.

2. *χ^2 approximations for the standardized and unstandardized cases.* For reference we note first the over-all χ^2 approximations. In the standardized case this is

$$\chi^2 = -\{n - \tfrac{1}{6}(2p+5)\}\log_e |R|, \tag{3}$$

with $\frac{1}{2}p(p-1)$ degrees of freedom. If V denotes the variance-covariance matrix only standardized in the more limited sense explained above, we have

$$\chi^2 = -\{n - \tfrac{1}{6}(2p+1+2/p)\}\log_e |V|, \tag{4}$$

with $\frac{1}{2}(p+2)(p-1)$ degrees of freedom. The approximation (3) follows from the moment formula

$$E\{|R|^t\} = \frac{\Gamma^{p-1}(\frac{1}{2}n)\prod_{i=1}^{p-1}\{\frac{1}{2}(n-i)+t\}}{\Gamma^{p-1}(\frac{1}{2}n+t)\prod_{i=1}^{p-1}\Gamma\frac{1}{2}(n-i)}. \tag{5}$$

This formula (originally given by Wilks) can be immediately deduced, for example, from the sampling distribution of the correlations r_{ij} in the null case, namely,

$$f(r_{ij}) = \frac{|R|^{\frac{1}{2}(n-p-1)}\Gamma^{p-1}(\frac{1}{2}n)}{\Pi^{\frac{1}{2}p(p-1)}\prod_{i=1}^{p-1}\Gamma\frac{1}{2}(n-1)}, \tag{6}$$

which itself follows from the Wishart distribution of variances and covariances.

The approximation (4) is most simply arrived at by combining the χ^2 approximation (3) with my χ^2 approximation for the homogeneity of the p variances, the latter χ^2 adding $p-1$ further degrees of freedom. More formally, from the known distribution of the latent roots ϕ_i of a variance matrix with true variance matrix the unit matrix, namely,

$$\frac{\pi^{\frac{1}{2}p}(\phi_1\phi_2\ldots\phi_p)^{\frac{1}{2}(n-p-1)}\exp\left(-\frac{1}{2}\sum_{i=1}^{p}\phi_i\right)\prod_{i=1}^{p}\prod_{j=i+1}^{p}(\phi_i-\phi_{i+j})}{2^{\frac{1}{2}pn}\prod_{i=0}^{p-1}\{\Gamma\frac{1}{2}(m-i)\,\Gamma\frac{1}{2}(p-i)\}}, \tag{7}$$

and that of $\phi_1+\phi_2+\ldots\phi_p$, namely,

$$\frac{(\phi_1+\phi_2+\ldots\phi_p)^{\frac{1}{2}pn-1}\exp\left(-\frac{1}{2}\sum_{i=1}^{p}\phi_i\right)}{2^{\frac{1}{2}pn}\Gamma(\frac{1}{2}pn)}, \tag{8}$$

we may obtain that of $u_i = p\phi_i \Big/ \sum_{j=1}^{p}\phi_j \quad (i = 1, 2, \ldots, p-1),$

as

$$\frac{\pi^{\frac{1}{2}p}(u_1u_2\ldots u_p)^{\frac{1}{2}(n-p-1)}\prod_{i=1}^{p}\prod_{j=i+1}^{p}(u_i-u_{i+j})}{p^{\frac{1}{2}np}\prod_{i=0}^{p-1}\{\Gamma\frac{1}{2}(n-i)\,\Gamma\frac{1}{2}(p-i)\}}, \tag{9}$$

whence

$$E\{|V|^t\} = \frac{p^{rp}\Gamma(\frac{1}{2}pn)\prod_{i=0}^{p-1}\Gamma\{\frac{1}{2}(n-i)+t\}}{\Gamma\{\frac{1}{2}p(n+2r)\}\prod_{i=0}^{p-1}\Gamma\frac{1}{2}(n-i)}, \tag{10}$$

and the χ^2 approximation (4) follows.

3. *The elimination of one heterogeneous component.* If the true correlations are not zero but all equal (to ρ, say), then the first true latent root is $1+(p-1)\rho$, and if this is removed the others are all homogeneous with the value $1-\rho$. It is proposed to examine this case in detail when n becomes large in order to examine the effect of standardization on the reduced χ^2.

We use $\delta\mathbf{V}$ to denote deviations from the true variance-covariance matrix

$$\mathbf{V} = (1-\rho)+\rho\boldsymbol{\xi}\boldsymbol{\xi}', \tag{11}$$

where for brevity the first term omits any symbol for the unit matrix, and $\boldsymbol{\xi}$ is a column vector with components unity (and $\boldsymbol{\xi}'$ its transpose). In the completely standardized case, the diagonal is standardized to unity, and no fluctuations in variance are possible, but it is convenient to handle both cases together at first. We now write, in the equation

$$|\mathbf{V}+\delta\mathbf{V}-\lambda| = 0, \tag{12}$$

$1-\lambda = \rho-x$, and we have exactly

$$\begin{vmatrix} \delta v_{11}-x & \delta v_{12}\dots \\ \delta v_{12} & \delta v_{22}-x\dots \\ \vdots & \end{vmatrix} + \rho\Sigma_1 \begin{vmatrix} 1 & \delta v_{12} & \delta v_{13} & \dots \\ 1 & \delta v_{22}-x & \delta v_{23} & \dots \\ \vdots & & & \end{vmatrix} = 0,$$

i.e.

$$\Delta_p + \rho\Sigma_1\Delta_1' = 0, \tag{13}$$

say, where Δ_p is $|\delta\mathbf{V}-x|$, Δ_1' denotes the modified determinant with its first column replaced by 1's, and Σ_1 denotes summation over permutations of the anomalous column 1. As we are assuming that ρ is not zero, the second term in (13) is of lower order than the first as n increases, and we may consider the asymptotically equivalent equation

$$\Sigma_1\Delta_1' = \Sigma_1\Delta_{p,1} - \Sigma_1\Sigma_2\delta v_{12}\Delta_{2,1}' = 0, \tag{14}$$

where $\Delta_{p,1}$ is a determinant like Δ_p with the first variable dropped, $\Delta_{2,1}'$ is like Δ_1' with the first variable dropped and 2 in place of 1, and Σ_2 denotes summation over the restricted permutations of the anomalous column 2 when the first variable has already been dropped. The terms of highest order in x in the determinant Δ_p itself are

$$(-x)^p - (-x)^{p-2}\Sigma_1\Sigma_2[(\delta v_{12})^2 - \delta v_{11}\delta v_{22}]\dots, \tag{15}$$

making use of the fact $\Sigma_1\delta v_{11} = 0$. Hence equation (14) becomes

$$\begin{aligned}&\Sigma_1\{(-x)^{p-1} - (-x)^{p-3}\Sigma_2\Sigma_3[(\delta v_{23})^2 - \delta v_{22}\delta v_{33}]\} \\ &\quad - \Sigma_1\Sigma_2\delta v_{12}\{(-x)^{p-2} + (-x)^{p-3}\Sigma_3\delta v_{33}\} + \Sigma_1\Sigma_2\Sigma_3\delta v_{12}\delta v_{23}(-x)^{p-3} + \dots = 0,\end{aligned}$$

where Σ_3 similarly denotes summation excluding the first two variables, or

$$\begin{aligned}&px^{p-1} + x^{p-2}(2\Sigma_1\Sigma_2\delta v_{12}) \\ &\quad + x^{p-3}\{2\Sigma_1\Sigma_2\Sigma_3\delta v_{12}(\delta v_{23}-\delta v_{33}) - (p-2)\Sigma_1\Sigma_2(\delta v_{23}^2 - \delta v_{22}\delta v_{33})\} + \dots = 0,\end{aligned} \tag{16}$$

from which the values of $\Sigma_1 x_1$ and $\Sigma_1\Sigma_2 x_1 x_2$, where x_i are the roots of (16), follow.

The asymptotic form of the reduced χ^2 (cf. formula (3) of my original paper) is, for large n,

$$-n\log\left[\frac{(p-1)^{p-1}\lambda_2\lambda_3\dots\lambda_p}{(\lambda_2+\lambda_3+\dots\lambda_p)^{p-1}}\right],$$

which, with $\lambda_i = 1-\rho+x_i$, becomes on expansion in terms of x_i up to the second degree,

$$\frac{n}{(1-\rho)^2}\left\{-\Sigma_1\Sigma_2 x_1 x_2+\frac{p-2}{2(p-1)}(\Sigma_1 x_1)^2\right\}. \tag{17}$$

Substituting from (16), we finally obtain

$$\frac{n}{(1-\rho)^2}\left\{\frac{(p-2)(p^2-p+2)}{p^2(p-1)}\Sigma_1\Sigma_2(\delta v_{12})^2+\frac{4(p-2)}{p^2(p-1)}\Sigma_1\Sigma_2\Sigma_3\Sigma_4\delta v_{12}\delta v_{34}\right.$$
$$\left.-\frac{2(p^2-3p+4)}{p^2(p-1)}\Sigma_1\Sigma_2\Sigma_3\delta v_{12}\delta v_{23}-\frac{p-2}{p}\Sigma_1\Sigma_2\delta v_{11}\delta v_{22}+\frac{2}{p}\Sigma_1\Sigma_2\Sigma_3\delta v_{12}\delta v_{33}\right\}. \tag{18a}$$

In the standardized case, when $\delta v_{ii}=0$, this reduces to

$$\frac{n}{(1-\rho)^2}\left\{\frac{(p-2)(p^2-p+2)}{p^2(p-1)}\Sigma_1\Sigma_2(\delta r_{12})^2+\frac{4(p-2)}{p^2(p-1)}\Sigma_1\Sigma_2\Sigma_3\Sigma_4\delta r_{12}\delta r_{34}\right.$$
$$\left.-\frac{2(p^2-3p+4)}{p^2(p-1)}\Sigma_1\Sigma_2\Sigma_3\delta r_{12}\delta r_{23}\right\}, \tag{18b}$$

where, since now v_{12} is a proper correlation, it is written r_{12}.

4. *The asymptotic expected value of* χ^2. To obtain the expected value of expressions like (18*a*) or (18*b*), we require various large-sample fluctuation formulae. The asymptotic distribution of the v_{ij} or r_{ij} being normal for large n, these may be obtained by standard methods. Two examples will be sufficient, one from each set.

(i) $$\operatorname{cov}(r_{12},r_{13})\sim\rho^2 E\left\{\left(\frac{\delta w_{12}}{\rho}-\tfrac{1}{2}\delta w_{11}-\tfrac{1}{2}\delta w_{22}\right)\left(\frac{\delta w_{13}}{\rho}-\tfrac{1}{2}\delta w_{11}-\tfrac{1}{2}\delta w_{33}\right)\right\},$$

where w_{12} denotes the completely unrestricted sample covariance (to distinguish it from v_{12} above), or since all true covariances are equal to ρ, this reduces to

$$\operatorname{cov}(w_{12},w_{13})-\rho\operatorname{cov}(w_{12},w_{11})-\rho\operatorname{cov}(w_{13},w_{22})+\tfrac{3}{4}\operatorname{cov}(w_{11},w_{22})+\tfrac{1}{4}\operatorname{var}(w_{11}).$$

The individual terms may be evaluated by the standard combinatorial rules, giving on reduction, under the above conditions,

$$n\operatorname{cov}(r_{12},r_{13})\sim\tfrac{1}{2}\rho(1-\rho)^2(2+3\rho). \tag{19}$$

Similarly, in addition to the well-known formula

$$n\operatorname{var}(r_{12})\sim(1-\rho^2)^2, \tag{20}$$

we obtain

$$n\operatorname{cov}(r_{12},r_{34})\sim 2\rho^2(1-\rho)^2. \tag{21}$$

(ii) $$\operatorname{cov}(v_{12},v_{11})\sim\rho E\left\{\left[\frac{\delta w_{12}}{\rho}-\frac{1}{p}(\delta w_{11}+\delta w_{22}+\ldots\delta w_{pp})\right]\left[\frac{p-1}{p}\delta w_{11}-\frac{1}{p}(\delta w_{22}+\ldots+\delta w_{pp})\right]\right\}$$
$$=-\frac{2}{p}\operatorname{cov}(w_{12},w_{11})+\frac{2}{p}\operatorname{cov}(w_{12},w_{33}),$$

the remaining terms having vanishing coefficients,

$$\sim-\frac{4\rho^2}{pn}(1-\rho). \tag{22}$$

Similarly
$$n \operatorname{var}(v_{12}) \sim (1-\rho)\left\{1+\rho+2\rho^2-2\rho^3-\frac{2\rho^2}{p}(3-\rho)\right\}, \tag{23}$$

$$n \operatorname{cov}(v_{12}, v_{13}) \sim \rho(1-\rho)\left\{1+2\rho-2\rho^2-\frac{2\rho}{p}(3-\rho)\right\}, \tag{24}$$

$$n \operatorname{cov}(v_{12}, v_{34}) \sim 2\rho^2(1-\rho)\left\{1-\rho-\frac{1}{p}(3-\rho)\right\}, \tag{25}$$

$$n \operatorname{cov}(v_{11}, v_{22}) \sim -\frac{2}{p}(1-\rho^2), \tag{26}$$

$$n \operatorname{cov}(v_{11}, v_{23}) \sim -\frac{4\rho}{p}(1-\rho). \tag{27}$$

Substituting these formulae in the average values of (18a) and (18b) we obtain after some algebra the expected values, in these two cases respectively,

$$\tfrac{1}{2}(p-2)(p+1) \tag{28a}$$

and
$$\tfrac{1}{2}(p-2)(p+1)-\frac{(p-1)(p-2)}{p}(1-\rho)^2. \tag{28b}$$

The first result merely checks with the result following from the more powerful and wider geometrical argument, but the second is of considerable interest. It indicates that the reduction in the total number of degrees of freedom, due to standardization, is only partially felt in the smaller roots, and disappears completely when the components eliminated take up the dominant part of the total variance. The degrees of freedom which I provisionally suggested for the reduced χ^2 in the standardized case are thus in general too low, and it is safer to allocate the number appropriate to the unstandardized case, namely,

$$\tfrac{1}{2}(p-k-1)(p-k+2),$$

when k components have been eliminated.

It is also evident from the expected value (28b) that the χ^2 approximation cannot strictly be valid in the standardized case except in the further limit $\rho \to 1$, corresponding to the residual components contributing negligibly to the total variance; this is in contrast with (28a), which, as anticipated, holds (for large n) for all $\rho \neq 0$. In this limiting case we might reasonably expect the reduced χ^2 for the standardized and unstandardized cases to coincide, but it is of interest further to demonstrate that the quadratic expression (18b) is then of the correct χ^2 form.

5. *The limiting quadratic form in the standardized case.* Let us denote the column *vector* of correlations $\delta r_{ij}\,(i,j = 1 \ldots p;\ j > i)$ by $\boldsymbol{\delta}$, and the corresponding variance matrix, with $\frac{1}{2}p(p-1)$ rows and columns, by $W(\rho)$. The quadratic form in (18b) may be denoted by $\boldsymbol{\delta}' A(\rho)\boldsymbol{\delta}$. Transforming to new standardized and uncorrelated variables

$$\boldsymbol{\eta} = W^{-\frac{1}{2}}(\rho)\boldsymbol{\delta}, \tag{29}$$

we have
$$\boldsymbol{\delta}' A(\rho)\boldsymbol{\delta} = \boldsymbol{\eta}' W^{\frac{1}{2}}(\rho) A(\rho) W^{\frac{1}{2}}(\rho)\boldsymbol{\eta}. \tag{30}$$

As $\rho \to 1$, we require to prove that

$$B = \lim_{\rho \to 1} W^{\frac{1}{2}}(\rho) A(\rho) W^{\frac{1}{2}}(\rho)$$

is a matrix of rank $\frac{1}{2}(p-2)(p+1)$ with latent roots 1 and 0. Since the factor $n/(1-\rho)^2$ in (18b) cancels with the factor $(1-\rho)^2/n$ in (19), (20) and (21), we can drop this factor, and then

put $\rho = 1$ throughout. We then find (using the notation W and A for these modified limiting matrices) that

$$W = \tfrac{3}{2} + \tfrac{5}{2}\xi\xi' - \tfrac{1}{2}P, \tag{31}$$

$$A = \frac{p-1}{p} - \frac{p^2-3p+4}{p^2(p-1)}\xi\xi' + \frac{1}{p}P, \tag{32}$$

where ξ denotes the 'unit' vector, and P denotes the symmetric matrix which has zero elements p_{rs} except when the suffices r and s, which are of the form $r \equiv (i,j)$, $s \equiv (k,l)$, are such that i, j, k and l are all different; in such cases p_{rs} is unity. For example, for $p = 5$, P is

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

when the order for r is taken as 12, 13, 14, 15, 23, 24, 25, 34, 35, 45.

It is evident that ξ is a latent vector of P with latent root $\frac{1}{2}(p-2)(p-3)$; also $\xi'\xi = \frac{1}{2}p(p-1)$. Suppose in consequence we assume that

$$P = \frac{(p-2)(p-3)}{p(p-1)}\xi\xi' + \Sigma\nu_s\xi_s\xi_s', \tag{33}$$

where ν_s are the remaining latent roots, and ξ_s the corresponding latent vectors (with $\xi_s'\xi_s = 1$). In view of the structure of A from (32), and the resolution of the unit matrix,

$$\frac{2}{p(p-1)}\xi\xi' + \Sigma\xi_s\xi_s',$$

we find

$$A = \Sigma\left(\frac{p-1}{p} + \frac{\nu_s}{p}\right)\xi_s\xi_s',$$

and further,

$$W^{\frac{1}{2}}AW^{\frac{1}{2}} = WA = \Sigma\left(\frac{3}{2} - \frac{\nu_s}{2}\right)\left(\frac{p-1}{p} + \frac{\nu_s}{p}\right)\xi_s\xi_s'. \tag{34}$$

Since $\frac{1}{2}p(p-1) - \frac{1}{2}(p-2)(p+1) = 1$, we require *all* the latent roots of the matrix in (34) to be unity, except for the zero root corresponding to $\xi\xi'$, which has dropped out. This requires that the values of all the remaining values ν_s satisfy the equation

$$(\tfrac{3}{4} - \tfrac{1}{2}x)\left(\frac{p-1}{p} + \frac{x}{p}\right) = 1,$$

or

$$(x+p-3)(x-1) = 0$$

or

$$x = +1 \quad \text{or} \quad -(p-3). \tag{35}$$

Since the sum of all the latent roots of P is 0, we find that its latent roots should be

$$\left.\begin{aligned} &\tfrac{1}{2}(p-2)(p-3)\ldots \text{once},\\ &-(p-3)\ldots \text{multiplicity } p-1,\\ &1\ldots \text{multiplicity } \tfrac{1}{2}p(p-3). \end{aligned}\right\}$$

This requirement represents a rather interesting property of the matrix P. I easily verified it for $p = 3$ and 4, but in the general case I am indebted to Dr Ledermann for a proof, which he has kindly summarized in the following appendix.

APPENDIX

BY W. LEDERMANN

In §5 of the preceding paper a matrix $P = (p_{rs})$ was introduced whose $\binom{p}{2} \equiv \frac{1}{2}p(p-1)$ rows and $\binom{p}{2}$ columns are specified by pairs of integers $r = (i,j), s = (k,l)$, where i,j,k,l run through the integers $1, 2, \ldots, p$, subject to $i<j$ and $k<l$. The elements of P are defined as follows:

$$p_{rs} = \begin{cases} 1, \text{ if } i,j,k,l \text{ are distinct,} \\ 0, \text{ in all other cases.} \end{cases}$$

We shall prove that the *matrix P has a simple spectral resolution and that,* if $p>3$, *its characteristic equation is*

$$|\lambda - P| = \{\lambda - \tfrac{1}{2}(p-2)(p-3)\}\{\lambda+p-3\}^{p-1}\{\lambda-1\}^{\frac{1}{2}p(p-3)}. \tag{1}$$

The first assertion, that is, the existence of a pure diagonal canonical form for P, follows at once from the fact that P is a real symmetric matrix.

Next, if it can be shown that P satisfies a polynomial equation

$$f(P) = 0, \tag{2}$$

then each latent root of P satisfies the scalar equation

$$f(z) = 0. \tag{3}$$

This follows from the fact that the canonical form of P also satisfies the matrix equation (2). It will presently be proved that P satisfies a cubic equation, whence we may infer that the three roots of this cubic are the latent roots of P with certain multiplicities.

From the definition of p_{rs} it is seen that each row and column contains exactly $\binom{p-2}{2}$ units and $\binom{p}{2} - \binom{p-2}{2}$ zeros. Since P is symmetric, the elements of P^2 are the scalar products of one row with another. Let the symbol $(\mathbf{i},\mathbf{j})'(\mathbf{u},\mathbf{v})$ denote the scalar product of the rows specified by the pairs (i,j) and (u,v) respectively. By counting the number of times that a unit occupies corresponding places in two given rows we find that

$$(\mathbf{i},\mathbf{j})'(\mathbf{u},\mathbf{v}) = \binom{p-2}{2} \quad \text{or} \quad \binom{p-3}{2} \quad \text{or} \quad \binom{p-4}{2},$$

according as the pairs (i,j) and (u,v) have two, one or no numbers in common. Corresponding to the three possible values of the elements of P^2 we write

$$P^2 = \binom{p-2}{2} I + \binom{p-3}{2} P_1 + \binom{p-4}{2} P_2,$$

where I is the unit matrix of order $\binom{p}{2}$ and P_1 (P_2) is a matrix with a unit element at the intersection of the row (i,j) and column (u,v) provided that these pairs of indices have exactly one (no) number in common, all other elements being zero.

Introducing the matrix (denoted by $\xi\xi'$ in §5 of the preceding paper)

$$E = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix},$$

we have $$P_1 = E - P - I, \quad P_2 = P.$$

Hence
$$P^2 = \binom{p-2}{2} I + \binom{p-3}{2} (E - P - I) + \binom{p-4}{2} P$$

$$= (p-3) I - (p-4) P + \binom{p-3}{2} E. \tag{4}$$

On multiplying by P and using the relation

$$EP = PE = \binom{p-2}{2} E,$$

we get
$$P^3 = (p-3) P - (p-4) P^2 + \binom{p-3}{2}\binom{p-2}{2} E, \tag{5}$$

whence, on eliminating E between (4) and (5),

$$P^3 - \left\{\binom{p-2}{2} - p + 4\right\} P^2 - \left\{\binom{p-2}{2}(p-4) + p - 3\right\} P + (p-3)\binom{p-2}{2} I = 0.$$

This is the cubic equation for P referred to above. Its roots are easily found to be

$$\binom{p-2}{2}, \quad -p+3, \quad 1. \tag{6}$$

It follows that
$$|\lambda - P| = \left\{\lambda - \binom{p-2}{2}\right\}^a \{\lambda + p - 3\}^b \{\lambda - 1\}^c,$$

where a, b, c are certain non-negative integers such that

$$a + b + c = \binom{p}{2}. \tag{7}$$

Also, since the trace of P is equal to the sum of the latent roots,

$$a\binom{p-2}{2} + b(-p+3) + c = 0. \tag{8}$$

Finally, we use the fact that the latent roots of P^2 are the squares of the quantities (6) with multiplicities a, b, c. Hence

$$a\binom{p-2}{2}^2 + b(-p+3)^2 + c = \binom{p}{2}\binom{p-2}{2}, \tag{9}$$

where the right-hand side is the trace of P^2. The numbers a, b, c can be determined from (7), (8) and (9), and it is found that

$$a = 1, \quad b = p - 1, \quad c = \tfrac{1}{2}p(p-3).$$

This establishes the equation (1).

REFERENCES

BARTLETT, M. S. (1938). *Proc. Camb. Phil. Soc.* **34**, 33.
BARTLETT, M. S. (1950). *Brit. J. Psychol.*, Statist. Sect. **3**, 77.
BARTLETT, M. S. (1951). *Brit. J. Psychol.*, Statist. Sect. **4**, 1.
BOX, G. E. P. (1949). *Biometrika*, **36**, 317.

THE JACOBIANS OF CERTAIN MATRIX TRANSFORMATIONS USEFUL IN MULTIVARIATE ANALYSIS

BASED ON LECTURES OF P. L. HSU AT THE UNIVERSITY OF NORTH CAROLINA, 1947

BY WALTER L. DEEMER AND INGRAM OLKIN, *University of North Carolina*

Editorial Note. The following paper was submitted by Prof. Hotelling in the summer of 1950 for publication in *Biometrika* with the accompanying Note from Col. Deemer and Mr Olkin:

'In 1947 Prof. P. L. Hsu gave courses in multivariate analysis at the University of North Carolina in which he developed new techniques for finding Jacobians of certain matrix transformations. Hsu returned to China at the end of that academic year, leaving as a record of this material only the notes of students in his classes.

'We (Deemer and Olkin) became interested in these matrix transformations in the course of our work in multivariate analysis. Since we did not take Hsu's courses, we used Ralph Bradley's notes as a basis for our studies.

'In the spring of 1948 there was a seminar in multivariate analysis under the direction of Prof. Harold Hotelling. At his request we prepared some lectures on matrix transformations. At the completion of the seminar, Prof. Hotelling and Prof. R. C. Bose suggested that in view of the importance of these techniques and their non-availability, we should prepare an expository paper giving a systematic development with all proofs given in detail.

'All the new ideas of importance in this paper are due to Hsu. Our contribution has been to organize the material in logical form, making all proofs complete with the necessary lemmas explicitly stated and proved.

'Efforts were made to communicate with Prof. Hsu in order that he could review this material before it was circulated. To date such efforts have failed.'

Since this contribution was received contact has been made with Prof. Hsu and the paper is now published with his approval. A suggestion which he made for improvement has been added as a Note on p. 361. However, in view of the liberty they have taken in the exposition of Hsu's ideas, the American authors would like it to be clear that they are to be held responsible for any errors.

FOREWORD

BY HAROLD HOTELLING

We are apparently at the beginning of a major development in the use of statistical procedures for joint treatment of a multiplicity of correlated variates. Many of the new methods depend ultimately on the distribution of the roots of certain determinantal equations. This distribution was published simultaneously in 1939 by P. L. Hsu (1939) and R. A. Fisher (1939) in the *Annals of Eugenics*, and, excepting for a constant multiplier, by S. N. Roy (1938–40) in *Sankhyā*. Hsu's derivation, which, like Roy's, is otherwise complete, demonstrates the correctness of his formula for this constant multiplier only for the case of three variates. Proof that the formula is correct for the general case has turned out to be unexpectedly difficult to develop. In seeking to close this gap while lecturing on advanced multivariate analysis at Chapel Hill, Hsu developed what appears to be an extraordinarily powerful method of dealing with a wide class of transformations of the elements of one matrix into those of another induced by a functional relation between the matrices. This technique not only leads to a solution of the original problem but appears to offer the possibility of a fresh attack on other problems of multivariate distribution theory as well. Heretofore it has

been recorded only in lecture notes made by students in the spring of 1947. Col. Deemer and Mr Olkin have not merely reproduced lecture notes but have reworked the material and improved it in certain essential respects.

1. Summary and introduction

In finding the distribution of multivariate statistics it is often desirable to make certain transformations. However, the calculation of the Jacobian is frequently rather difficult. We have therefore assembled in this paper the Jacobians of many of the most used transformations.

Certain general methods are applied in the calculation of the Jacobians; these methods will be found useful in the calculation of other Jacobians not given here. Three techniques of particular value for non-linear transformations are:

(*a*) Taking differentials to get a linear equation in the differentials, e.g. Theorem 4·1.

(*b*) Transformations to new variables so that the Jacobian can be evaluated as the product of other Jacobians which are more easily calculated, e.g. Theorem 4·2.

(*c*) Introduction of new matrices such that the Jacobian of the original transformation is the discriminant of the bilinear form in the variables of the new matrices, e.g. Lemma 4·3 and Theorem 4·3.

2. Notation

Because the shape and symmetry of matrices are important in finding Jacobians, we shall consistently use the following notation:

(2·1) Lower-case letters will denote scalars.

(2·2) As is customary, a prime will denote the transpose of a matrix.

(2·3) Lower-case letters with a superior bar will denote column vectors, e.g. $\bar{u}' = (u_1, \ldots, u_p)$.

(2·4) Capital letters will denote matrices. A system of superior symbols will be used as follows:

(*a*) $\bar{A}$ is a symmetric matrix of elements a_{ij}.

(*b*) $\hat{A}$ is a skew-symmetric matrix of elements a_{ij} above the main diagonal and $a_{ji} = -a_{ij}$ below the main diagonal and zeros in the main diagonal.

(*c*) $\vec{A}$ is a triangular matrix with zeros above the main diagonal, e.g.

$$\begin{bmatrix} a_{11} & \cdots & \cdots & \cdot \\ a_{12} & a_{22} & \cdots & \cdot \\ \cdots & \cdots & \cdots & \cdot \\ a_{1p} & a_{2p} & \cdots & a_{pp} \end{bmatrix}.$$

(*d*) A matrix not in classes (*a*), (*b*), (*c*) will be denoted by a capital letter with no superior symbol.

(*e*) (dX) is the matrix of the same shape as X, whose elements are the differentials of the elements of X. This will be denoted by X^*.

(2·5) Diagonal matrices will be denoted by D with a subscript showing the elements, e.g. D_r is the diagonal matrix with elements $r_1, r_2, \ldots, r_p$.

(2·6) Orthogonal matrices will be denoted by Δ and Γ.

(2·7) $A: p \times q$ means the matrix A has p rows and q columns; the same notation will be used for matrices with superior symbols.

(2·8) Since it is the absolute value of the Jacobian that is wanted, it is to be understood that all determinants are evaluated up to the sign only.

(2·9) Let X, Y be two matrices which satisfy an equation $Y = f(X)$. Let X and Y have the same number n of independent elements, and denote them by $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, respectively. Then the Jacobian, $\dfrac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}$, will be denoted by $D(X; Y)$. For purposes of calculation it is sometimes desirable to express this in terms of the elements of X and Y, for example, $D(X; Y)$ might be written $D(x_{ij}, x_{kl}; y_{ij}, y_{kl})$.

More generally, if $X_1, \ldots, X_k$, $Y_1, \ldots, Y_m$ satisfy the equations $Y_i = f_i(x_1, \ldots, x_k)$ $(i = 1, \ldots, m)$ and $(X_1, \ldots, X_k)$ and $(Y_1, \ldots, Y_m)$ have the independent elements $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, respectively, then $D(X_1, \ldots, X_k; Y_1, \ldots, Y_m)$ denotes

$$\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}.$$

(2·10) The following abbreviations will be used: p.d. for positive definite; n.s. for non-singular; J for Jacobian; tr for trace; I_p for the identity matrix of order p.

3. THE JACOBIANS OF LINEAR TRANSFORMATIONS

We state here a number of theorems on the J of linear transformations, some without proof. When the J does not immediately follow from previous theorems we will give the proof.

THEOREM 3·1 *a*. The J of the transformation $Y = aX$ $(Y, X: p \times q)$, where a is a constant scalar $\neq 0$, is $D(X; Y) = a^{pq}$.

Proof. $y_{ij} = ax_{ij}$ and $\partial y_{ij}/\partial x_{ij} = a$. Therefore the scheme of coefficients is diagonal with pq a's in the main diagonal. Q.E.D.

THEOREM 3·1 *b*. The J of the transformation $\overline{Y} = a\overline{X}$. $(\overline{Y}, \overline{X}: p \times p)$, where a is a constant scalar $\neq 0$, is $D(\overline{X}; \overline{Y}) = a^{\frac{1}{2}p(p+1)}$.

THEOREM 3·2. The J of the linear transformation $\overline{y} = A\overline{x}$ $(\overline{y}, \overline{x}: p \times 1;\ A: p \times p)$ is $D(\overline{x}; \overline{y}) = |A|$.

Proof. $y_i = \sum_{k=1}^{p} a_{ik}x_k$ and $\partial y_i/\partial x_j = a_{ij}$, therefore $J = D(\overline{x}; \overline{y}) = |A|$. Q.E.D.

COROLLARY 3·2. The J of the transformation $\overline{x} = A^{-1}\overline{y}$ $(\overline{x}, \overline{y}: p \times 1;\ A: p \times p)$ is $D(\overline{y}; \overline{x}) = |A|^{-1}$.

This follows from Theorem 3·2, since $|A^{-1}| = |A|^{-1}$.

THEOREM 3·3. The J of the transformation

$$\overline{y}' = \overline{x}'A' (\overline{y}, \overline{x}: p \times 1;\ A: p \times p) \quad \text{is} \quad D(\overline{x}; \overline{y}) = |A|.$$

THEOREM 3·4. The J of the transformation

$$Y = AX (Y, X: p \times q;\ A: p \times p) \quad \text{is} \quad D(X; Y) = |A|^q.$$

Proof. This follows from Theorem 3·2 and 5 B. 4†, since the transformation of each column of Y is independent of the others and there are q such columns of Y, the J of each column transformation being $|A|$.

THEOREM 3·5. The J of $Y = XA$ $(Y, X: p \times q;\ A: q \times q)$ is $D(X; Y) = |A|^p$.

Proof. Write $Y' = A'X'$ and the result follows directly from Theorem 3·4.

† References of this form are to the appendix to the present paper.

THEOREM 3·6. The J of the transformation

$$Y = AXB (Y, X: p \times q;\ A: p \times p;\ B: q \times q) \quad \text{is} \quad D(X;\ Y) = |A|^q |B|^p.$$

Proof. Let $Z = AX$ and then $Y = ZB$. $D(X; Z) = |A|^q$ and $D(Z; Y) = |B|^p$. Using 5 B. 2 the theorem follows.

LEMMA 3·7. Let $$Y = A_n A_{n-1} \dots A_1 X A_1' \dots A_{n-1}' A_n',$$

then $$D(X;\ Y) = D(X;\ Y_1)\, D(Y_1;\ Y_2) \dots D(Y_{n-1};\ Y),$$

where $$Y_i = A_i Y_{i-1} A_i' \quad (i = 1, \dots, n); \qquad Y_0 \equiv X, \quad Y_n \equiv Y.$$

Proof. Follows directly from 5 B. 2.

THEOREM 3·7. The J of the transformation

$$\overline{Y} = B \overline{X} B' (\overline{Y}, B, \overline{X}: p \times p) \quad \text{is} \quad D(\overline{X};\ \overline{Y}) = |B|^{p+1}.$$

Proof. By 5 A. 5 we can write $B = F_m F_{m-1} \dots F_1$, where some of the F's are of the form 5 A. 1 and the rest are of the form 5 A. 2. Then

$$\overline{Y} = F_m \dots F_1 \overline{X} F_1' \dots F_m'. \tag{3·7·1}$$

By Lemma 3·7: $$D(\overline{X};\ \overline{Y}) = D(\overline{X};\ \overline{Y}_1)\, D(\overline{Y}_1;\ \overline{Y}_2) \dots D(\overline{Y}_{m-1};\ \overline{Y}_m),$$

where $$\overline{Y}_\nu = F_\nu \overline{Y}_{\nu-1} F_\nu' \quad (\nu = 1, \dots, m); \qquad \overline{Y}_0 \equiv \overline{X}, \quad \overline{Y}_m \equiv \overline{Y}.$$

We show that $$D(\overline{Y}_{\nu-1};\ \overline{Y}_\nu) = |F_\nu|^{p+1} \quad (\nu = 1, \dots, m),$$

and hence it will follow that

$$D(\overline{X};\ \overline{Y}) = |F_m|^{p+1} \dots |F_1|^{p+1} = |F_m \dots F_1|^{p+1} = |B|^{p+1}.$$

(1) Let G be any of the F's which are of the form 5 A. 1; the transformation $\overline{Y}_\nu = G \overline{Y}_{\nu-1} G'$ implies $$y_{ii} = a^2 x_{ii}; \qquad y_{ij} = a x_{ij} \quad (i \neq j); \qquad y_{jk} = x_{jk} \quad (j, k \neq i).$$

The determinant of the partial derivatives is therefore diagonal with $p-1$ elements a, one element a^2, and the remaining elements 1. Therefore $D(\overline{Y}_{\nu-1};\ \overline{Y}_\nu) = a^{p+1} = |G|^{p+1}$.

(2) Let H be any of the F's which are of the form 5 A. 2; the transformation $\overline{Y}_\nu = H \overline{Y}_{\nu-1} H'$ implies

$$y_{ii} = x_{ii} + 2a x_{ij} + a^2 x_{jj},$$

$$y_{ki} = y_{ik} = x_{ik} + a x_{jk} \quad (k \neq i),$$

$$y_{jk} = x_{jk} \quad (j, k \neq i).$$

The determinant of partial derivatives will have 1's in the main diagonal and 0's below the main diagonal. Hence $$D(\overline{Y}_{\nu-1};\ \overline{Y}_\nu) = 1 = |H|^{p+1}.$$ Q.E.D.

COROLLARY 3·7. The J of the transformation

$$\overline{Y} = A^{-1} \overline{X} A'^{-1} \ (\overline{Y}, \overline{X}, A: p \times p) \quad \text{is} \quad D(\overline{X};\ \overline{Y}) = |A|^{-(p+1)}.$$

THEOREM 3·8. The J of the transformation $\bar{Y} = \bar{A} \bar{X}$ $(\bar{Y}, \bar{A}, \bar{X}: p \times p)$ is $D(\bar{X};\ \bar{Y}) = \prod_{i=1}^{p} a_{ii}^i$.

Proof. $$y_{ij} = \sum_{\nu=i}^{j} a_{\nu j} x_{i\nu} \quad (i \leqslant j).$$

The elements on one side of the main diagonal of the determinant of partial derivatives are zero. The main diagonal contains a_{11} from $\partial y_{11}/\partial x_{11}$; it contains a_{22} twice: once from $\partial y_{12}/\partial x_{12}$, and once from $\partial y_{22}/\partial x_{22}$; it contains a_{ii} i times: once from each of $\partial y_{ji}/\partial x_{ji}$ $(j = 1, \dots, i)$.

Therefore the determinant is $\prod_{i=1}^{p} a_{ii}^i$. Q.E.D.

THEOREM 3·9. The J of the transformation

$$\bar{Y} = \bar{A}\bar{X}' + \bar{X}\bar{A}' \quad (\bar{Y}, \bar{A}, \bar{X}: p \times p) \quad \text{is} \quad D(\bar{X}; \bar{Y}) = 2^p \prod_{i=1}^{p} a_{ii}^{p-i+1}.$$

Proof. Pre- and post-multiply by $\bar{A}^{-1}$ and $\bar{A}'^{-1}$ respectively:

$$\bar{A}^{-1}\bar{Y}\bar{A}'^{-1} = \bar{X}'\bar{A}'^{-1} + \bar{A}^{-1}\bar{X}. \tag{3·9·1}$$

Let $\bar{Z} = \bar{A}^{-1}\bar{Y}\bar{A}'^{-1}$ and $\bar{U} = \bar{A}^{-1}\bar{X}$. Then

$$\bar{Z} = \bar{U}' + \bar{U}, \tag{3·9·2}$$

$$D(\bar{X}; \bar{Y}) = D(\bar{X}; \bar{U})\, D(\bar{U}; \bar{Z})\, D(\bar{Z}; \bar{Y}) \tag{3·9·3}$$

by 5 B. 2;

$$D(\bar{X}; \bar{U}) = \prod_{i=1}^{p} a_{ii}^{-i} \tag{3·9·4}$$

by Theorem 3·8;

$$D(\bar{Z}; \bar{Y}) = |\bar{A}|^{p+1} = \prod_{i=1}^{p} a_{ii}^{p+1} \tag{3·9·5}$$

by Corollary 3·7 and 5 B. 1;

$$z_{ii} = 2u_{ii} \quad \text{and} \quad z_{ij} = u_{ij} \tag{3·9·6}$$

from 3·9·2. The determinant of partial derivatives is therefore diagonal. It contains p elements equal to 2 and the remainder 1's. Therefore

$$D(\bar{U}; \bar{Z}) = 2^p. \tag{3·9·7}$$

Substituting in (3·9·3) $\quad D(\bar{X}; \bar{Y}) = 2^p \prod_{i=1}^{p} a_{ii}^{p-i+1}$. Q.E.D.

THEOREM 3·10. The J of the transformation

$$\hat{Y} = G\hat{X}G' \quad (\hat{Y}, \hat{X}, G: p \times p) \quad \text{is} \quad D(\hat{X}; \hat{Y}) = |G|^{p-1}.$$

Proof. The proof follows along the same lines as the proof of Theorem 3·7 with the necessary changes due to $\hat{X}$ and $\hat{Y}$ being skew symmetric.

4. NON-LINEAR TRANSFORMATIONS

The general method for finding the J of a non-linear transformation is to take the differential of both sides and then to find the J of the resulting linear transformation in the differentials (see 5 B. 3).

THEOREM 4·1. The J of the transformation $\bar{V} = \bar{T}\bar{T}'$ is $D(\bar{T}; \bar{V}) = 2^p \prod_{i=1}^{p} t_{ii}^{p-i+1}$.

Proof. Taking differentials of $\bar{V} = \bar{T}\bar{T}'$,

$$d\bar{V} = \bar{T}(d\bar{T})' + (d\bar{T})\bar{T}'. \tag{4·1·1}$$

Let $\bar{V}^* = d\bar{V}$; $\bar{T}^* = d\bar{T}$, then

$$\bar{V}^* = \bar{T}\bar{T}^{*\prime} + \bar{T}^*\bar{T}'. \tag{4·1·2}$$

This is the transformation of Theorem 3·9, and hence the theorem follows. Q.E.D.

LEMMA 4·2. Let $\bar{A}$, $\bar{B}$ be p.d., $p \times p$ matrices. Let $|\bar{A} - \phi\bar{B}| = 0$ have p distinct roots $\phi_1 > \phi_2 > \ldots > \phi_p > 0$, and let D_ϕ be the diagonal matrix with elements $\phi_1, \ldots, \phi_p$. Then there exists a unique matrix $W = (w_{ij})$: $p \times p$, where $w_{1i} > 0$ $(i = 1, \ldots, p)$ such that $\bar{A} = WD_\phi W'$ and $\bar{B} = WW'$.

THEOREM 4·2.† The J of the transformation $\bar{A} = WD_\phi W'$; $\bar{B} = WW'$ (all matrices $p \times p$) is

$$D(W, \phi; \bar{A}, \bar{B}) = |W|^{p+2} 2^p \prod_{i<j}^{p} (\phi_i - \phi_j).$$

Proof. Taking differentials:

$$d\bar{A} = (dW) D_\phi W' + W(dD_\phi) W' + WD_\phi (dW)', \tag{4·2·1}$$

$$d\bar{B} = W(dW)' + (dW) W'. \tag{4·2·2}$$

Let $$\bar{A}^* = d\bar{A}; \quad \bar{B}^* = d\bar{B}; \quad W^* = dW; \quad D_\eta = dD_\phi,$$

then

$$\bar{A}^* = W^* D_\phi W' + WD_\eta W' + WD_\phi W^{*\prime}, \tag{4·2·3}$$

$$\bar{B}^* = WW^{*\prime} + W^* W'. \tag{4·2·4}$$

By 5 B. 3: $J = D(W^*, \eta; \bar{A}^*, \bar{B}^*)$. Pre- and post-multiply (4·2·3) and (4·2·4) by W^{-1} and W'^{-1} respectively:

$$W^{-1}\bar{A}^* W'^{-1} = W^{-1}W^* D_\phi + D_\eta + D_\phi W^{*\prime} W'^{-1}, \tag{4·2·5}$$

$$W^{-1}\bar{B}^* W'^{-1} = W^{*\prime} W'^{-1} + W^{-1} W^*. \tag{4·2·6}$$

Let

$$\bar{U} = W^{-1}\bar{A}^* W'^{-1}, \tag{4·2·7}$$

and hence

$$\bar{A}^* = W\bar{U}W';$$

$$\bar{V} = W^{-1}\bar{B}^* W'^{-1}, \tag{4·2·8}$$

and hence

$$\bar{B}^* = W\bar{V}W';$$

$$R = W^{-1}W^*. \tag{4·2·9}$$

Substituting in (4·2·5) and (4·2·6):

$$\bar{U} = RD_\phi + D_\phi R' + D_\eta, \tag{4·2·10}$$

$$\bar{V} = R + R'; \tag{4·2·11}$$

$$D(W^*, \eta; \bar{A}^*, \bar{B}^*) = D(W^*, \eta; R, \eta) D(R, \eta; \bar{U}, \bar{V}) D(\bar{U}, \bar{V}; \bar{A}^*, \bar{B}^*) \tag{4·2·12}$$

using 5 B. 2;

$$D(W^*, \eta; R, \eta) = |W|^{-p} \tag{4·2·13}$$

from (4·2·9) using Theorem 3·4 and 5 B. 4;

$$D(\bar{U}, \bar{V}; \bar{A}^*, \bar{B}^*) = D(\bar{U}; \bar{A}^*) D(\bar{V}; \bar{B}^*) \tag{4·2·14}$$

by 5 B. 4;

$$D(\bar{U}; \bar{A}^*) = D(\bar{V}; \bar{B}^*) = |W|^{p+1} \tag{4·2·15}$$

by Theorem 3·7. Hence

$$D(\bar{U}, \bar{V}; \bar{A}^*, \bar{B}^*) = |W|^{2(p+1)}. \tag{4·2·16}$$

To find $D(R, \eta; \bar{U}, \bar{V})$ we use (4·2·10) and (4·2·11):

$$u_{ii} = 2r_{ii}\phi_i + \eta_i, \qquad v_{ii} = 2r_{ii},$$

$$u_{ij} = r_{ij}\phi_j + r_{ji}\phi_i, \quad v_{ij} = r_{ij} + r_{ji} \quad (i \neq j).$$

† For an application see Hsu (1939).

The scheme of partial derivatives will then be as follows:

	η_i $(i=1,\dots,p)$	r_{ii} $(i=1,\dots,p)$	r_{ij} $(i>j)$	r_{ij} $(i<j)$
u_{ii} $(i=1,\dots,p)$	I	L	0	0
v_{ii} $(i=1,\dots,p)$	0	$2I$	0	0
u_{ij} $(i<j)$	0	0	D	C
v_{ij} $(i<j)$	0	0	I	I

where C is the diagonal matrix with elements $\phi_2, \phi_3, \dots, \phi_p; \phi_3, \dots, \phi_p; \dots; \phi_{p-1}, \phi_p; \phi_p$, and D is the diagonal matrix with elements $\phi_1, \dots, \phi_1$ ($p-1$ terms), $\phi_2, \dots, \phi_2$ ($p-2$ terms), ..., ϕ_{p-1}.

The matrix L does not affect the value of the determinant, which is equal to

$$2^p \,|D-C| = 2^p \prod_{i<j}^{p} (\phi_i - \phi_j)$$

by applying 5 A. 7 twice. Hence

$$J = 2^p \,|W|^{p+2} \prod_{i<j}^{p} (\phi_i - \phi_j). \quad \text{Q.E.D.}$$

LEMMA 4·3. Let $X_1 = (x_{ij}^{(1)})$, $(p_1 \times q_1)$ and $X_2 = (x_{ij}^{(2)})$, $(p_2 \times q_1)$ be two matrices containing a total of N independent variables $(N \leqslant p_1 q_1 + p_2 q_1)$ and $Y_1 = (y_{ij}^{(1)})$, $(p_1 \times q_1)$, $Y_2 = (y_{ij}^{(2)})$, $(p_2 \times q_1)$ be two other matrices containing the same number of variables.

Let Ξ: $N \times 1$ be the vector formed from all N x's by using the elements in the first column of X_1 as the first p_1 elements of Ξ, the p_1 elements in the second column of X_1 as the next p_1 elements of Ξ, etc., through the columns of X_2.

Let H: $N \times 1$ be the vector formed from the columns of Y_1 and Y_2 in the same way that Ξ was formed from X_1 and X_2. Thus

$$\Xi' = (x_{11}^{(1)} x_{21}^{(1)} \dots x_{p_1 q_1}^{(1)} x_{11}^{(2)} x_{21}^{(2)} \dots x_{p_2 q_1}^{(2)}),$$

$$\mathrm{H}' = (y_{11}^{(1)} y_{21}^{(1)} \dots y_{p_1 q_1}^{(1)} y_{11}^{(2)} y_{21}^{(2)} \dots y_{p_2 q_1}^{(2)}).$$

Let $\Xi = A\mathrm{H}$ (A: $N \times N$), where A is a matrix of constants, be the connecting equations between the variables $x_{ij}^{(\nu)}$ and $y_{ij}^{(\nu)}$ ($\nu = 1, 2$).

Let $R_1 = (r_{ij}^{(1)})$, $(q_1 \times p_1)$ and $R_2 = (r_{ij}^{(2)})$, $(q_1 \times p_2)$ be two matrices with a total of N independent variables. Then $D(Y_1, Y_2; X_1, X_2)$ is equal to the discriminant of $\mathrm{tr}\,(R_1 X_1 + R_2 X_2)$, where $\mathrm{tr}\,(R_1 X_1 + R_2 X_2)$ is to be considered as a bilinear form in the variables of R_1, R_2 and Y_1, Y_2.

Proof. Let $\mu(N \times 1)$ be the vector formed from the elements R_1 and R_2 by taking the rows of R_1 in turn and then the rows of R_2 in turn:

$$\mu' = (r_{11}^{(1)} r_{12}^{(1)} \dots r_{p_1 q_1}^{(1)} r_{11}^{(2)} r_{12}^{(2)} \dots r_{p_2 q_1}^{(2)}),$$

$$\mathrm{tr}\,(R_1 X_1 + R_2 X_2) = \sum_{\nu=1}^{2} \sum_{i=1}^{q_1} \sum_{\sigma=1}^{p_\nu} r_{i\sigma}^{(\nu)} x_{\sigma i}^{(\nu)} = \mu' \Xi = \mu' A \mathrm{H}.$$

The discriminant of this bilinear form is $|A|$.

The J, $D(Y_1, Y_2; X_1, X_2)$ of the original transformation $\Xi = A\mathrm{H}$ is $|A|$ (by Theorem 3·2). Q.E.D.

Introduction to Theorem 4·3

Let $\overline{V} = \begin{bmatrix} \overline{V}_{11} & V_{12} \\ V_{21} & \overline{V}_{22} \end{bmatrix}$ be p.d. ($\overline{V}_{11}$: $p \times p$; V_{12}: $p \times q$; $\overline{V}_{22}$: $q \times q$) ($p \leqslant q$). There exists a n.s. matrix G of the form $\begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}$ (G_1: $p \times p$; G_2: $q \times q$) such that

$$G\overline{V}G' = \begin{bmatrix} I_p & D_r & 0 \\ D_r & & I_q \\ 0 & & \end{bmatrix}, \tag{4·3·2}$$

where D_r: $p \times p$ has elements $r_1, \ldots, r_p$; $r_1^2 \ldots r_p^2$ are the roots of the determinantal equation $|V_{12}\overline{V}_{22}^{-1}V_{21} - \theta\overline{V}_{11}| = 0$ (see 5 A. 21).

$$\overline{V} = G^{-1}\begin{bmatrix} I_p & D_r & 0 \\ D_r & & I_q \\ 0 & & \end{bmatrix}G'^{-1} \tag{4·3·3}$$

from (4·3·2). Let

$$G^{-1} = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix},$$

then

$$\overline{V} = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} I_p & D_r & 0 \\ D_r & & I_q \\ 0 & & \end{bmatrix}\begin{bmatrix} P' & 0 \\ 0 & Q' \end{bmatrix}. \tag{4·3·4}$$

This is a transformation from the v's to the p's, q's and r's. $\overline{V}$ contains $\frac{1}{2}(p+q)(p+q+1)$ variables, and the right side of (4·3·4) contains p^2+q^2+p variables. Let $\delta = q-p$, then $[(p^2+q^2+p) - \frac{1}{2}(p+q)(p+q+1)] = \frac{1}{2}\delta(\delta-1)$, and the number of variables on each side will be equal if $p = q$ or $p = q-1$.

We now partition Q and rewrite (4·3·4):

$$\overline{V} = \begin{bmatrix} P & 0 & 0 \\ 0 & Q_{11} & Q_{12} \\ 0 & Q_{21} & Q_{22} \end{bmatrix}\begin{bmatrix} I_p & D_r & 0 \\ D_r & I_p & 0 \\ 0 & 0 & I_\delta \end{bmatrix}\begin{bmatrix} P' & 0 & 0 \\ 0 & Q'_{11} & Q'_{21} \\ 0 & Q'_{12} & Q'_{22} \end{bmatrix}, \tag{4·3·5}$$

(Q_{11}: $p \times p$; Q_{12}: $p \times \delta$; Q_{22}: $\delta \times \delta$).

Post-multiplying Q_{12} and Q_{22} by an orthogonal matrix Δ will not change the value of this product, i.e.

$$\overline{V} = \begin{bmatrix} P & 0 & 0 \\ 0 & Q_{11} & Q_{12}\Delta \\ 0 & Q_{21} & Q_{22}\Delta \end{bmatrix}\begin{bmatrix} I_p & D_r & 0 \\ D_r & I_p & 0 \\ 0 & 0 & I_\delta \end{bmatrix}\begin{bmatrix} P' & 0 & 0 \\ 0 & Q'_{11} & Q'_{21} \\ 0 & \Delta'Q'_{12} & \Delta'Q'_{22} \end{bmatrix},$$

as may be verified by multiplying out this expression. By 5 A. 10 we can choose this Δ so that $Q_{22}\Delta$ is triangular with 0's below the main diagonal. When we do this we reduce the number of variables on the right by $\frac{1}{2}\delta(\delta-1)$ and thus have the same number on each side. The transformation is unique if $r_1 > \ldots > r_p > 0$; $p_{ij} > 0$ ($j = 1, \ldots, p$). We now find its J after the number of variables is equalized.

Hsu has used this transformation in finding the distribution of the canonical correlation coefficient (see Hotelling, 1936) when the population values are not zero. Hsu's work on this has apparently not been published.

THEOREM 4·3. Let $$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & T' \end{bmatrix}$$

(Q_{11}: $p \times p$; Q_{12}: $p \times \delta$; T: $\delta \times \delta$) ($p \leqslant q$), $\delta = q-p$; $t_{ii} > 0$ ($i = 1, \ldots, \delta$).

Let $$\Phi = \begin{bmatrix} I_p & D_r & 0 \\ D_r & & I_q \\ 0 & & \end{bmatrix},$$

(D_r: $p \times p$). Then the J of the transformation

(4·3·6) $$\overline{V} = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \Phi \begin{bmatrix} P' & 0 \\ 0 & Q' \end{bmatrix}$$

is $$D(P, Q, r; \overline{V}) = 2^{p+q} |P|^{q+1} |Q|^{p+1} \left(\prod_{i=1}^{s} t_{ii}^{i-1} \right) \left(\prod_{i=1}^{p} r_i^q \right) \left(\prod_{i<j}^{p} (r_i^2 - r_j^2) \right).$$

When $q = p$ or $q = p+1$, the terms in t do not appear.

Proof. Take differentials of (4·3·6) and let $\overline{V}^* = d\overline{V}$, $P^* = dP$, $Q_{11}^* = dQ_{11}$, $Q_{12}^* = dQ_{12}$, $Q_{21}^* = dQ_{21}$, $\overline{T}^* = d\overline{T}$, $D_\eta = dD_r$, and let

$$Q^* = dQ = \begin{bmatrix} Q_{11}^* & Q_{12}^* \\ Q_{21}^* & \overline{T}^{*\prime} \end{bmatrix};$$

then

(4·3·7) $$\overline{V}^* = \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix} \Phi \begin{bmatrix} P' & 0 \\ 0 & Q' \end{bmatrix} + \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \Phi \begin{bmatrix} P^* & 0 \\ 0 & Q^* \end{bmatrix}' + \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} 0 & D_\eta & 0 \\ D_\eta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P' & 0 \\ 0 & Q' \end{bmatrix}.$$

$J = D(P^*, Q^*, \eta; \overline{V}^*)$ by 5 B. 3. Pre-multiply (4·3·7) by $\begin{bmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix}$ and post-multiply by $\begin{bmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix}'$ and introduce new variables:

(4·3·8) $$\overline{X} = \begin{bmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} \overline{V}^* \begin{bmatrix} P^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix}',$$

(4·3·9) $$A = P^{-1} P^*,$$

(4·3·10) $$\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = Q^{-1} \begin{bmatrix} Q_{11}^* \\ Q_{21}^* \end{bmatrix}.$$

To simplify the notation, we write

(4·3·11) $$\begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix} = Q^{-1} \begin{bmatrix} Q_{12}^* \\ \overline{T}^{*\prime} \end{bmatrix},$$

but we do not consider B_{12} and B_{22} as new variables, since $Q^{-1}\overline{T}^{*\prime}$ is not triangular, and hence the number of variables in B_{22} would be greater than the number of variables in $\overline{T}^{*\prime}$. We then get from (4·3·7):

(4·3·12) $$\overline{X} = \begin{bmatrix} A & 0 & 0 \\ 0 & B_{11} & B_{12} \\ 0 & B_{21} & B_{22} \end{bmatrix} \Phi + \Phi \begin{bmatrix} A' & 0 & 0 \\ 0 & B_{11}' & B_{21}' \\ 0 & B_{12}' & B_{22}' \end{bmatrix} + \begin{bmatrix} 0 & D_\eta & 0 \\ D_\eta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} A + A' & AD_r + D_r B_{11}' + D_\eta & D_r B_{21}' \\ (AD_r + D_r B_{11}' + D_\eta)' & B_{11} + B_{11}' & B_{12} + B_{21}' \\ (D_r B_{21}')' & (B_{12} + B_{21}')' & B_{22} + B_{22}' \end{bmatrix}.$$

By 5 B. 2

(A) $$J = D(P^*, Q^*, \eta; \overline{V}^*)$$
$$= D(P^*, Q^*, \eta; A, B_{11}, B_{21}, Q_{12}^*, \overline{T}^*, \eta)\, D(A, B_{11}, B_{21}, Q_{12}^*, \overline{T}^*, \eta; \overline{X})\, D(\overline{X}; \overline{V}^*)$$
$$= D_1 D_2 D_3.$$

By 5 B. 4 $$D_1 = D(P^*; A)\, D(Q_{11}^*, Q_{21}^*; B_{11}, B_{21})$$
$$= |P|^{-p} |Q|^{-p}$$

(from (4·3·9), (4·3·10), using Theorem 3·4). From (4·3·8), using Theorem 3·7,

(B) $$D_3 = D(\overline{X}; \overline{V}^*) = |P|^{p+q+1} |Q|^{p+q+1}.$$

To compute D_2 we partition $\bar{X}$ and hence can equate the submatrices of $\bar{X}$ to the submatrices on the right of (4·3·12):

$$\bar{X} = \begin{bmatrix} \bar{X}_{11} & X_{12} & X_{13} \\ X'_{12} & \bar{X}_{22} & X_{23} \\ X'_{13} & X'_{23} & \bar{X}_{33} \end{bmatrix} \quad (\bar{X}_{11}, X_{12}, \bar{X}_{22}: p \times p;\ X_{13}, X_{23}: p \times \delta;\ \bar{X}_{33}: \delta \times \delta).$$

We then have

$$\text{(4·3·13)} \qquad \bar{X}_{11} = A + A';\quad X_{12} = AD_\tau + D_\tau B'_{11} + D_\eta;\quad \bar{X}_{22} = B_{11} + B'_{11},$$

$$\text{(4·3·14)} \qquad X_{13} = D_\tau B'_{21};\quad X_{23} = B_{12} + B'_{21};\quad \bar{X}_{33} = B_{22} + B'_{22}.$$

The transformations in (4·3·13) and (4·3·14) are independent, so we can write

$$\text{(C)} \qquad D_2 = D(A, B_{11}, \eta; \bar{X}_{11}, X_{12}, \bar{X}_{22})\, D(B_{21}, Q^*_{12}, \hat{T}^*; X_{13}, X_{23}, \bar{X}_{33})$$
$$= D_4 D_5.$$

Computation of D_4. For convenience we will relabel the elements of $\bar{X}$:

$$\text{(4·3·15)} \qquad \bar{X}_{11} = (x_{ij});\quad X_{12} = (y_{ij});\quad \bar{X}_{22} = (z_{ij}),$$

all $i, j = 1, \ldots, p$, and we denote the elements of the other matrices as follows:

$$A = (a_{ij});\quad B'_{11} = (b_{ij})\quad (i, j = 1, \ldots, p);\qquad \eta = (\eta_1 \ldots \eta_p).$$

Using (4·3·13), (4·3·14) and (4·3·15) we get

$$\text{(4·3·16)} \qquad x_{ii} = 2a_{ii};\quad y_{ii} = r_i a_{ii} + r_i b_{ii} + \eta_i;\quad z_{ii} = 2b_{ii},$$

$$\text{(4·3·17)} \qquad x_{ij} = a_{ij} + a_{ji}\quad (i<j);\qquad y_{ij} = r_j a_{ij} + r_i b_{ij}\quad (i \neq j);\qquad z_{ij} = b_{ij} + b_{ji}\quad (i<j).$$

The transformations in (4·3·16) and (4·3·17) are independent, hence

$$D_4 = D(A, B_{11}, \eta; \bar{X}_{11}, X_{12}, \bar{X}_{22}) = D(a_{ii}, b_{ii}, \eta_i; x_{ii}, y_{ii}, z_{ii})\, D(a_{ij}, b_{ij}; x_{ij}, y_{ij}, z_{ij}).$$

For $D(a_{ii}, b_{ii}, \eta_i; x_{ii}, y_{ii}, z_{ii})$ we have the scheme of coefficients:

	a_{ii}	b_{ii}	η_i	
x_{ii}	$2I$	0	0	
z_{ii}	0	$2I$	0	$= 2^{2p},$
y_{ii}	L	M	I	

the values L and M not affecting the determinant. To compute $D(a_{ij}, b_{ij}; x_{ij}, y_{ij}, z_{ij})$ we rewrite in (4·3·17):

$$y_{ij} = r_j a_{ij} + r_i b_{ij}\quad (i<j = 1, \ldots, p),$$
$$y_{ji} = r_i a_{ji} + r_j b_{ji}\quad (i<j = 1, \ldots, p).$$

We then have the scheme of coefficients ($i<j$ in all cases)

	a_{ij}	b_{ji}	a_{ji}	b_{ij}
x_{ij}	I	0	I	0
z_{ij}	0	I	0	I
y_{ji}	0	C	D	0
y_{ij}	C	0	0	D

$$= \left| \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} - \begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix} \right| = |D|\,|D - CD^{-1}C| = |D^2 - C^2|,$$

using 5 A. 6 and 5 A. 7 twice. C is a diagonal matrix with elements

$$r_2, r_3, \ldots, r_p;\ r_3, \ldots, r_p;\ \ldots;\ r_{p-1}, r_p;\ r_p;$$

D is a diagonal matrix with elements

$$r_1, \ldots, r_1\,(p-1 \text{ terms}), r_2, \ldots, r_2\,(p-2 \text{ terms}), \ldots, r_{p-1}.$$

Therefore
$$D^2 - C^2 = \prod_{i<j}^{p} (r_i^2 - r_j^2),$$

and hence

$$\text{(D)} \qquad D_4 = D(A, B_{11}, \eta; \bar{X}_{11}, X_{12}, \bar{X}_{22}) = 2^{2p} \prod_{i<j}^{p} (r_i^2 - r_j^2).$$

Computation of $D(B_{21}, Q_{12}^*, \bar{T}^*; X_{13}, X_{23}, \bar{X}_{33}) = D_5$. The connecting equations are given in (4·3·14). B_{22} is independent of the other variables, but recall that the original variables are Q_{12}^* and Q_{22}^* and that the relation (4·3·11) does not define B_{12} and B_{22} as new variables.

The scheme of coefficients, from (4·3·14), is

$$\text{(4·3·18)} \qquad \begin{array}{c|c|cc} & B_{21} & Q_{12}^* & \bar{T}^* \\ \hline X_{13} & F & 0 & 0 \\ \hline X_{23} & L & \multicolumn{2}{c}{G} \\ \bar{X}_{33} & & & \end{array} = |F|\,|G|,$$

so the value of L does not affect the value of the determinant. $|F|$ is the J of the transformation $X_{13} = D_r B_{21}'$ $(X_{13}, B_{21}'\colon p \times \delta;\ D_r\colon p \times p)$ which by Theorem 3·4 is $|D_r|^\delta = \left(\prod_{i=1}^{p} r_i\right)^\delta$. Since the value of B_{21} does not affect G, we can, for determining G, write $X_{23} = B_{12}$ and $\bar{X}_{33} = B_{22} + B_{22}'$. Therefore

$$\text{(E)} \qquad D_5 = \left(\prod_{i=1}^{p} r_i\right)^\delta D_6; \quad D_6 = D(Q_{12}^*, \bar{T}^*; X_{23}, \bar{X}_{33}).$$

Computation of D_6. The connecting equations between B_{12}, B_{22}; Q_{12}^*, $\bar{T}^*$; X_{23}, $\bar{X}_{33}$ are

$$\begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix} = Q^{-1} \begin{bmatrix} Q_{12}^* \\ \bar{T}^{*\prime} \end{bmatrix},$$

$$X_{23} = B_{12},$$

$$\bar{X}_{33} = B_{22} + B_{22}'.$$

As previously stated, B_{12} and B_{22} cannot be considered as new variables. We therefore do not have any direct equations between the variables of X_{23}, $\bar{X}_{33}$ and Q_{12}^*, $\bar{T}^*$. It is clear, however, that the elements of X_{23}, $\bar{X}_{33}$ are linear functions of the elements of Q_{12}^*, $\bar{T}^*$, and hence we can apply Lemma 4·3.

Let $W = (w_{ij})\colon (\delta \times p)$ and $\bar{S}\colon (\delta \times \delta)$, where $\bar{S} = (\epsilon_{ij} s_{ij})$ $(i, j = 1, \ldots, \delta)$; $\epsilon_{ij} = 1$ if $i = j$; $\epsilon_{ij} = \frac{1}{2}$ if $i \neq j$.

Let
$$\begin{aligned} \operatorname{tr} H &= \operatorname{tr}(W X_{23} + \bar{S}\bar{X}_{33}) \\ &= \operatorname{tr}(W B_{12} + \bar{S} B_{22} + \bar{S} B_{22}') = \operatorname{tr}(W B_{12} + 2\bar{S} B_{22}) \text{ (by 5 A. 12)}; \\ &= \operatorname{tr}\left((W\ 2\bar{S}) \begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix}\right) \\ &= \operatorname{tr}\left((W\ 2\bar{S})\, K \begin{bmatrix} Q_{12}^* \\ \bar{T}^{*\prime} \end{bmatrix}\right), \end{aligned}$$

where $K = Q^{-1}$.

Writing $Q_{12}^* = (\beta_{ij})$ $(i = 1, \dots, p;\ j = 1, \dots, \delta)$, $\bar{T}^* = (\tau_{ij})$ $(i \leqslant j = 1, \dots, \delta)$:

$$(4\cdot3\cdot19) \qquad \operatorname{tr} H = \operatorname{tr} \left[\begin{array}{ccc|ccc} w_{11} & \dots & w_{1p} & 2s_{11} & s_{12} & \dots & s_{1\delta} \\ \dots & \dots & \dots & s_{12} & 2s_{22} & \dots & s_{2\delta} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ w_{\delta 1} & \dots & w_{\delta p} & s_{1\delta} & \dots & \dots & 2s_{\delta\delta} \end{array}\right] K \begin{bmatrix} \beta_{11} & \dots & \beta_{1\delta} \\ \dots & \dots & \dots \\ \beta_{p1} & \dots & \beta_{p\delta} \\ \tau_{11} & \dots & \tau_{1\delta} \\ \dots & \dots & \dots \\ 0 & \dots & \tau_{\delta\delta} \end{bmatrix}.$$

H is $\delta \times \delta$ and $\operatorname{tr} H = \sum_{j=1}^{\delta} h_{jj}$. h_{jj} is found by multiplying the jth row of $(W\ 2\bar{S})$ into K and then multiplying the resulting row vector into the jth column of $\begin{bmatrix} Q_{12}^* \\ \bar{T}^{*\prime} \end{bmatrix}$. Therefore

$$\operatorname{tr} H = \sum_{j=1}^{\delta} (w_{j1} \dots w_{jp} s_{j1} \dots 2s_{jj} \dots s_{j\delta})\, K \begin{bmatrix} \beta_{1j} \\ \vdots \\ \beta_{pj} \\ \tau_{1j} \\ \vdots \\ \tau_{jj} \end{bmatrix}.$$

We now partition K so that the partition is a function of j:

$$(4\cdot3\cdot20) \qquad Q^{-1} = K = \begin{bmatrix} K_{11}^{(j)} & K_{12}^{(j)} \\ K_{21}^{(j)} & K_{22}^{(j)} \end{bmatrix} \quad (1 \leqslant j \leqslant \delta),$$

$$(K_{11}^{(j)}\colon (p+j) \times (p+j);\ K_{12}^{(j)}\colon (p+j) \times (\delta - j);\ K_{22}^{(j)}\colon (\delta - j) \times (\delta - j)).$$

Now we can write

$$(4\cdot3\cdot21) \qquad \operatorname{tr} H = \sum_{j=1}^{\delta} \left\{ (w_{j1} \dots w_{jp} s_{j1} \dots 2s_{jj}) K_{11}^{(j)} \begin{bmatrix} \beta_{1j} \\ \vdots \\ \beta_{pj} \\ \tau_{1j} \\ \vdots \\ \tau_{jj} \end{bmatrix} + (s_{j,j+1} \dots s_{j\delta}) K_{21}^{(j)} \begin{bmatrix} \beta_{1j} \\ \vdots \\ \beta_{pj} \\ \tau_{1j} \\ \vdots \\ \tau_{jj} \end{bmatrix} \right\}.$$

By Lemma 4·3, $D(Q_{12}^*, \bar{T}^*; X_{23}, \bar{X}_{33})$ is the determinant of this bilinear form. Writing the terms for $j = 1, 2, \dots, \delta$ in order we find that the discriminant has 0's above the main diagonal, and the product of the terms in the main diagonal is

$$(4\cdot3\cdot22) \qquad 2\,|K_{11}^{(1)}|\,2\,|K_{11}^{(2)}| \dots 2\,|K_{11}^{(\delta)}| = 2^{\delta} \prod_{j=1}^{\delta} |K_{11}^{(j)}|;$$

$$|K_{11}^{(j)}| = \begin{vmatrix} k_{11} & \dots & k_{1,p+j} \\ \dots & \dots & \dots \\ k_{p+j,1} & \dots & k_{p+j,p+j} \end{vmatrix},$$

where the elements k_{ih} are the elements of Q^{-1}. Let A_{ih} be an element of the adjoint of Q. Then $k_{ih} = \dfrac{1}{|Q|} A_{hi}$. Therefore

$$|K_{11}^{(j)}| = \frac{1}{|Q|^{p+j}} \begin{vmatrix} A_{11} & \dots & A_{1,p+j} \\ \dots & \dots & \dots \\ A_{p+j,1} & \dots & A_{p+j,p+j} \end{vmatrix} = \frac{1}{|Q|} \begin{vmatrix} q_{p+j+1,p+j+1} & \dots & q_{p+j+1,p+\delta} \\ \dots & \dots & \dots \\ q_{p+\delta,p+j+1} & \dots & q_{p+\delta,p+\delta} \end{vmatrix},$$

by 5 A. 11. But $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & \bar{T}' \end{bmatrix}$ $(Q_{11}: p \times p;\ Q_{12}: p \times \delta;\ \bar{T}': \delta \times \delta)$.

Therefore $$q_{p+k+1,\,p+h+1} = t_{k+1,\,h+1}.$$

Therefore

$$|K_{11}^{(Q)}| = \frac{1}{|Q|}\begin{vmatrix} t_{j+1,j+1} & \cdots & t_{j+1,\delta} \\ \cdots & \cdots & \cdots \\ 0 & \cdots & t_{\delta\delta} \end{vmatrix} = \frac{1}{|Q|}\, t_{j+1,j+1} \cdots t_{\delta\delta}. \tag{4·3·23}$$

On substituting this result in (4·3·22) we find

$$D_6 = D(Q_{12}^*, \bar{T}^*; X_{23}, \bar{X}_{33}) = 2^\delta \prod_{j=1}^{\delta} |K_{11}^{(Q)}| = 2^\delta\, |Q|^{-\delta}\, t_{22} t_{33}^2 \cdots t_{\delta\delta}^{\delta-1}. \tag{F}$$

We now consolidate the results from equations (A) through (F):

$$J = 2^{p+q}\, |P|^{q+1}\, |Q|^{p+1} \left(\prod_{i=2}^{\delta} t_{ii}^{i-1}\right)\left(\prod_{i=1}^{p} r_i\right)^{\delta} \prod_{i<j}^{p} (r_i^2 - r_j^2). \quad \text{Q.E.D.}$$

THEOREM 4·4. The J of the transformation

$$\bar{U} = |\bar{V}|\, \bar{V}^{-1}, \tag{4·4·1}$$

where $\bar{V}: p \times p$ is p.d., is $D(\bar{V}; \bar{U}) = (p-1)\, |\bar{V}|^{\frac{1}{2}(p+1)(p-2)}$.

Proof. Take differentials of (4·4·1) and let $\bar{U}^* = d\bar{U}$, $\bar{V}^* = d\bar{V}$; then

$$\bar{U}^* = |\bar{V}|\, \mathrm{tr}\,(\bar{V}^{-1}\bar{V}^*)\, \bar{V}^{-1} - |\bar{V}|\, (\bar{V}^{-1}\bar{V}^*\bar{V}^{-1}), \tag{4·4·2}$$

using 5 A. 13, 5 A. 14 and 5 A. 15.

By 5 B. 3, $$J = D(\bar{V}^*; \bar{U}^*).$$

$\bar{V}$, and hence $\bar{V}^{-1}$, is p.d., so by 5 A. 16 we can factor $\bar{V}^{-1}$ into $\bar{T}\bar{T}'$. Substituting this for $\bar{V}^{-1}$ in (4·4·2) and pre- and post-multiplying the result by $\bar{T}'^{-1}$ and $\bar{T}^{-1}$ respectively,

$$\bar{T}'^{-1}\bar{U}^*\bar{T}^{-1} = |\bar{V}|\, \mathrm{tr}\,(\bar{T}'\bar{V}^*\bar{T})\, I - |\bar{V}|\, (\bar{T}'\bar{V}^*\bar{T}). \tag{4·4·3}$$

Let

$$\begin{cases} \bar{Z} = \bar{T}'^{-1}\bar{U}^*\bar{T}^{-1} & (\bar{Z}: p \times p), \\ \bar{W} = \bar{T}'\bar{V}^*\bar{T} & (\bar{W}: p \times p). \end{cases} \tag{4·4·4}$$

By Theorem 3·7 and Corollary 3·7

$$\begin{cases} D(\bar{U}^*; \bar{Z}) = |\bar{T}|^{-(p+1)}, \\ D(\bar{V}^*; \bar{W}) = |\bar{T}|^{p+1}. \end{cases} \tag{4·4·5}$$

Putting (4·4·4) in (4·4·3)

$$\bar{Z} = |\bar{V}|\, (\mathrm{tr}\,\bar{W})\, I - |\bar{V}|\, \bar{W}. \tag{4·4·6}$$

Then
$$\begin{aligned} D(\bar{V}^*; \bar{U}^*) &= D(\bar{V}^*; \bar{W})\, D(\bar{W}; \bar{Z})\, D(\bar{Z}; \bar{U}^*) \quad \text{(by 5 B. 2)} \\ &= |\bar{T}|^{p+1}\, D(\bar{W}; \bar{Z})\, |\bar{T}|^{p+1} \\ &= |\bar{V}|^{-(p+1)}\, D(\bar{W}; \bar{Z}). \end{aligned}$$

Computation of $D(\bar{W}; \bar{Z})$. Let

$$\bar{S} = |\bar{V}|\, \bar{W} \quad (\bar{S}: p \times p). \tag{4·4·7}$$

By Theorem 3·1 *b*: $$D(\bar{W}; \bar{S}) = |\bar{V}|^{\frac{1}{2}p(p+1)}.$$

Substitute (4·4·7) in (4·4·6):

$$\bar{Z} = (\mathrm{tr}\,\bar{S})\, I - \bar{S}, \tag{4·4·8}$$

whence we find

$$z_{ii} = \sum_{\nu=1}^{p} s_{\nu\nu} - s_{ii} = \sum_{\substack{\nu=1 \\ \nu \neq i}}^{p} s_{\nu\nu}. \tag{4·4·9}$$

Therefore $$\frac{\partial z_{ii}}{\partial s_{ii}} = 0; \qquad \frac{\partial z_{ii}}{\partial s_{jj}} = 1 \quad (i \neq j); \qquad \frac{\partial z_{ii}}{\partial s_{ij}} = 0 \quad (i \neq j).$$

$$z_{ij} = -s_{ij} \quad (i<j),$$

and therefore $$\frac{\partial z_{ij}}{\partial s_{ij}} = -1.$$

The scheme of coefficients is

(4·4·10)

	s_{ii}	s_{ij} $(i<j)$
z_{ii}	H	0
z_{ij} $(i<j)$	0	$-I$

the determinant of which is

$$H = \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix} = D(\bar{S}; \bar{Z}). \tag{4·4·11}$$

Let $\bar{h}$ be a column vector, $p \times 1$, consisting entirely of 1's. Then $\bar{h}\bar{h}'$ is a $p \times p$ matrix consisting entirely of 1's, and $D(\bar{S}; \bar{Z}) = |\, \bar{h}\bar{h}' - I_p \,|$. Using 5 A. 8 we get

$$D(\bar{S}; \bar{Z}) = |\, 1 - \bar{h}'\bar{h} \,| = p - 1.$$

Consolidating the results

$$\begin{aligned} J = D(\bar{V}^*; \bar{U}^*) &= D(\bar{V}^*; \bar{W})\, D(\bar{W}; \bar{S})\, D(\bar{S}; \bar{Z})\, D(\bar{Z}; \bar{U}^*) \\ &= |\, \bar{V} \,|^{-(p+1)} |\, \bar{V} \,|^{\frac{1}{2}p(p+1)} (p-1) \\ &= (p-1) |\, \bar{V} \,|^{\frac{1}{2}(p+1)(p-2)}. \quad \text{Q.E.D.} \end{aligned}$$

COROLLARY 4·4. The J of the inverse of the transformation of Theorem 4·4 is

$$D(\bar{U}; \bar{V}) = \frac{1}{p-1} |\, \bar{U} \,|^{-(p+1)(p-2)/[2(p-1)]}.$$

Proof.

$$|\, \bar{U} \,| = |\, \bar{V} \,|^{p} |\, \bar{V} \,|^{-1} = |\, \bar{V} \,|^{p-1},$$

and therefore

$$|\, \bar{V} \,| = |\, \bar{U} \,|^{1/(p-1)}.$$

$$D(\bar{U}; \bar{V}) = \frac{1}{D(\bar{V}; \bar{U})} = \frac{1}{p-1} |\, \bar{V} \,|^{-\frac{1}{2}(p+1)(p-2)}$$

$$= \frac{1}{p-1} |\, \bar{U} \,|^{-(p+1)(p-2)/[2(p-1)]}. \quad \text{Q.E.D.}$$

THEOREM 4·5.† The J of the transformation $U = \hat{T}[2(\hat{X}+I)^{-1} - I]$ is

$$D(\hat{T}, \hat{X}; U) = 2^{\frac{1}{2}p(p-1)} \left(\prod_{i=1}^{p-1} t_{ii}^{p-i} \right) |\, \hat{X} + I \,|^{-(p-1)} \quad (\hat{X}, \hat{T}, U: p \times p;\ U \text{ n.s.}).$$

This transformation is unique if $|\, \hat{T}^{-1}U + I \,| \neq 0$ (see 5 A. 18).

† This transformation has been used by P. L. Hsu in unpublished work on the distribution of roots of determinantal equations under certain non-null hypotheses.

Proof. Since $\hat{X}$ is skew symmetric it does not have a characteristic root -1, so $|\hat{X}+I| \neq 0$. By 5 A. 17 and 5 A. 19 we can write $UU' = \tilde{T}\tilde{T}'$. Therefore $\tilde{T}^{-1}UU'\tilde{T}'^{-1} = I$, and hence $\Gamma = \tilde{T}^{-1}U$ is orthogonal. Then

$$U = \tilde{T}\Gamma = \tilde{T}[2(\hat{X}+I)^{-1} - I] \quad \text{(by 5 A. 18).} \tag{4·5·1}$$

Take differentials of (4.5.1) and put $U^* = dU$; $\tilde{T}^* = d\tilde{T}$; $\hat{X}^* = d\hat{X}$. Then using 5 A. 20 we get

$$U^* = \tilde{T}^*\Gamma - \tfrac{1}{2}\tilde{T}(\Gamma+I)\hat{X}^*(\Gamma+I), \tag{4.5.2}$$

and by 5 B. 3

$$J = D(\tilde{T}^*, \hat{X}^*; U^*). \tag{4·5·3}$$

Let

$$\begin{cases} W = \tilde{T}^{-1}U^*\Gamma', \\ \tilde{R} = \tilde{T}^{-1}\tilde{T}^*, \\ \hat{Y} = (\Gamma+I)\hat{X}^*(\Gamma'+I). \end{cases} \tag{4·5·4}$$

Pre- and post-multiply (4·5·2) by $\tilde{T}^{-1}$ and Γ' respectively and substitute from (4·5·4). Then

$$W = \tilde{R} - \tfrac{1}{2}\hat{Y}. \tag{4·5·5}$$

By 5 B. 2:

$$D(\tilde{T}^*, \hat{X}^*; U^*) = D(\tilde{T}^*, \hat{X}^*; \tilde{R}, \hat{Y})\, D(\tilde{R}, \hat{Y}; W)\, D(W; U^*), \tag{4·5·6}$$

which by 5 B. 4 becomes

$$= D(\tilde{T}^*; \tilde{R})\, D(\hat{X}^*; \hat{Y})\, D(\tilde{R}, \hat{Y}; W)\, D(W; U^*).$$

From (4·5·4), using Theorems 3·6, 3·8 and 3·10,

$$D(W; U^*) = |\tilde{T}|^p \quad \text{(since } |\Gamma| = 1\text{)},$$

$$D(\tilde{T}^*; \tilde{R}) = \prod_{i=1}^{p} t_{ii}^{-i},$$

$$D(\hat{X}^*; \hat{Y}) = |\Gamma+I|^{p-1}.$$

From (4·5·5)

$$w_{ii} = r_{ii}; \quad w_{ij} = -\tfrac{1}{2}y_{ij} \quad (i<j);$$

$$w_{ji} = r_{ij} + \tfrac{1}{2}y_{ij} \quad (i<j).$$

The scheme of coefficients is

		r_{ii}	y_{ij} $(i<j)$	r_{ij} $(i<j)$
w_{ii}		I	0	0
w_{ij}	$(i<j)$	0	$-\tfrac{1}{2}I$	0
w_{ji}	$(i<j)$	0	$\tfrac{1}{2}I$	I

Therefore

$$D(\tilde{R}, \hat{Y}; W) = 2^{-\frac{1}{2}p(p-1)}.$$

Hence, consolidating the results,

$$J = 2^{-\frac{1}{2}p(p-1)}\left(\prod_{i=1}^{p-1} t_{ii}^{p-i}\right)|\Gamma+I|^{p-1}.$$

By 5 A. 19,

$$|\Gamma+I|^{p-1} = 2^{p(p-1)}|\hat{X}+I|^{-(p-1)},$$

and we have

$$J = 2^{\frac{1}{2}p(p-1)}\left(\prod_{i=1}^{p-1} t_{ii}^{p-i}\right)|\hat{X}+I|^{-(p-1)}. \quad \text{Q.E.D.}$$

THEOREM 4·6. The J of the transformation

$$\bar{S} = [2(\hat{X}+I)^{-1} - I]\, D_\theta [2(\hat{X}+I)^{-1} - I]'$$

is
$$D(\hat{X}, \theta; \bar{S}) = 2^{\frac{1}{2}p(p-1)} |\hat{X}+I|^{-(p-1)} \prod_{i<j}^{p} (\theta_i - \theta_j) \quad (\bar{S}, D_\theta, \hat{X}: p \times p).$$

Proof. By 5 A. 19 we can write

$$\bar{S} = \Gamma D_\theta \Gamma'. \tag{4·6·1}$$

Let

$$\begin{cases} \bar{S}^* = d\bar{S}, \\ \hat{X}^* = d\hat{X}, \\ D_\eta = dD_\theta. \end{cases} \tag{4·6·2}$$

Take differentials of (4·6·1) using 5 A. 20 and substitute from (4·6·2):

$$\bar{S}^* = -\tfrac{1}{2}(\Gamma+I)\hat{X}^*(\Gamma+I) D_\theta \Gamma' + \tfrac{1}{2}\Gamma D_\theta (\Gamma'+I)\hat{X}^*(\Gamma'+I) + \Gamma D_\eta \Gamma'. \tag{4·6·3}$$

(Note that $\hat{X}' = -\hat{X}$.) By 5 B. 3

$$J = D(\hat{X}^*, \eta; \bar{S}^*). \tag{4·6·4}$$

Let

$$\begin{cases} \bar{V} = \Gamma' \bar{S}^* \Gamma, \\ \hat{Y} = (\Gamma'+I)\hat{X}^*(\Gamma+I). \end{cases} \tag{4·6·5}$$

Pre- and post-multiply (4·6·3) by Γ' and Γ respectively and substitute from (4·6·5):

$$\bar{V} = -\tfrac{1}{2}\hat{Y} D_\theta + \tfrac{1}{2} D_\theta \hat{Y} + D_\eta, \tag{4·6·6}$$

whence from 5 B. 4

$$D(\hat{X}^*, \eta; \bar{S}^*) = D(\hat{X}^*, \eta; \hat{Y}, \eta)\, D(\hat{Y}, \eta; \bar{V})\, D(\bar{V}; \bar{S}^*). \tag{4·6·7}$$

From (4·6·5), using Theorems 3·7 and 3·10,

$$\begin{cases} D(\bar{V}; \bar{S}^*) = 1, \\ D(\hat{X}^*; \hat{Y}) = |\Gamma + I|^{p-1}. \end{cases} \tag{4·6·8}$$

From (4·6·6): $v_{ii} = \eta_i$; $v_{ij} = \tfrac{1}{2}\theta_i y_{ij} - \tfrac{1}{2}\theta_j y_{ij} = \tfrac{1}{2}(\theta_i - \theta_j) y_{ij}$ $(i<j)$.

The scheme of coefficients is

$$\begin{array}{c|cc} & \eta_i & y_{ij} \\ \hline v_{ii} & I & 0 \\ v_{ij} & 0 & H \end{array},$$

where H is diagonal, with elements

$$\tfrac{1}{2}(\theta_1 - \theta_2),\ \ldots,\ \tfrac{1}{2}(\theta_1 - \theta_p),\ \tfrac{1}{2}(\theta_2 - \theta_3),\ \ldots,\ \tfrac{1}{2}(\theta_{p-1} - \theta_p),$$

a total of $\frac{1}{2}p(p-1)$ terms. Therefore

$$D(\hat{Y}, \eta; \bar{V}) = 2^{-\frac{1}{2}p(p-1)} \prod_{i<j}^{p} (\theta_i - \theta_j), \tag{4·6·9}$$

and

$$J = |\Gamma + I|^{p-1} 2^{-\frac{1}{2}p(p-1)} \prod_{i<j}^{p} (\theta_i - \theta_j); \tag{4·6·10}$$

whence, using 5 A. 18, $J = 2^{\frac{1}{2}p(p-1)} |\hat{X}+I|^{-(p-1)} \prod_{i<j}^{p} (\theta_i - \theta_j)$. Q.E.D.

Note. The following comments relative to the uniqueness of the transformation of Theorem 4·6 have been made by P. L. Hsu in a communication to the Editor:

'THEOREM. The equation

$$\bar{S} = [2(I+\hat{X})^{-1} - I]\, D_\theta [2(I+\hat{X})^{-1} - I]'$$

establishes a (1, 1)-correspondence between $\bar{S}$ and $\hat{X}$, provided that

(i) $\bar{S}$ does not belong to a certain set of symmetric matrices which constitutes a set of measure zero (in the $\frac{1}{2}p(p+1)$-dimensional space), and

(ii) the elements (1, 2), (1, 3), ..., (1, p) of the matrix $(I+\hat{X})^{-1}$ are all positive.

Remark on the above theorem. The condition (ii) on $\hat{X}$ may be put in a simpler form as follows: if we write

$$\hat{X} = \begin{pmatrix} 0 & -\bar{x} \\ \bar{x}' & \hat{Y} \end{pmatrix},$$

then the row formed of the elements (1, 2), (1, 3), ..., (1, p) of $(I+\hat{X})^{-1}$ is equal to

$$\frac{\bar{x}(I+\hat{Y})^{-1}}{1+\bar{x}(I+\hat{Y})^{-1}\bar{x}'}.$$

Thus an equivalent to the condition (ii) is the following: $(\text{ii})_1$ every element of the row $\bar{x}(I+\hat{Y})^{-1}$ is positive.

As to the proof of the above theorem, which is not very short, I intend to give it in a paper to be published in one of the journals issued by the Chinese Academy of Sciences.'

5. APPENDIX

A. *Certain matrix theorems*

The following definitions and results from the theory of matrices and determinants are stated for reference purposes.

5 A. 1. Multiplication of the ith $\begin{matrix}\text{row}\\ \text{column}\end{matrix}$ of a matrix by a constant $a \neq 0$ may be accomplished by $\begin{matrix}\text{pre-}\\ \text{post-}\end{matrix}$multiplication by a diagonal matrix D, with a in the ith position and 1's elsewhere. $|D| = a$; D^{-1} is a matrix of the same type as D (see Frazer, Duncan & Collar (1938), § 3·11).

5 A. 2. To add $(a) \times \begin{matrix}(j\text{th row})\\ (i\text{th column})\end{matrix}$ to the $\begin{matrix}i\text{th row}\\ j\text{th column}\end{matrix}$, $a \neq 0$, $\begin{matrix}\text{pre-}\\ \text{post-}\end{matrix}$multiply by a matrix E containing 1's in the main diagonal, a in the ith row and jth column, and 0's elsewhere. $|E| = 1$; E^{-1} is a matrix of the same type as E (see Frazer *et al.* (1938), § 3·11).

Any combination or succession of a finite number of operations of types 5 A. 1 and 5 A. 2 is called an *elementary transformation.*

5 A. 3. The rank of a matrix is unchanged by an elementary transformation (see Bôcher (1938), § 19).

5 A. 4. *Canonical forms.* A $p \times p$ matrix A of rank r can be reduced by an elementary transformation to a canonical form, i.e. to a diagonal matrix with 1's in the first r positions and 0's elsewhere. The reduction to a diagonal matrix of the characteristic roots can be effected by an orthogonal transformation (see Bôcher (1938), p. 280).

If the matrix A is $p \times q$, $p < q$, of rank r, there exist two orthogonal matrices $\Gamma: p \times p$ and $\Delta: q \times q$ such that $\Gamma' A \Delta = (\Lambda 0)$, where

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix} \quad (p \times p)$$

and $\lambda_1^2 \dots \lambda_p^2$ are the characteristic roots of the matrix AA'. Notice that $p - r$ of the λ_i are zero (see Hsu, *Quart. J. Math.*, Oxford Series, **17** (1946), 162).

Proof. AA' is non-negative (see Albert (1946), p. 64), and hence its roots are non-negative and can be written $\lambda_1^2 \dots \lambda_r^2$. The rank $(A) = r$ and hence there exists an orthogonal matrix Γ such that

$$\Gamma' AA' \Gamma = \begin{bmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_p^2 \end{bmatrix} = \Lambda^2$$

of the theorem (see Cramér, 1946, 11·9). Since we can rearrange the position of the λ's in Λ by an orthogonal transformation we can assume

$$\Lambda^2 = \begin{bmatrix} \lambda_1^2 & & & & 0 \\ & \ddots & & & \\ & & \lambda_r^2 & & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix} = \begin{bmatrix} \Lambda_1^2 & 0 \\ 0 & 0 \end{bmatrix}$$

and $|\Lambda_1| \neq 0$.

Write $\Gamma = (\Gamma_1 \Gamma_2)$, where $\Gamma_1: p \times r$; $\Gamma_2: p \times p - r$. Then

$$\Lambda^2 = \Gamma' AA' \Gamma = \begin{bmatrix} \Gamma_1' \\ \Gamma_2' \end{bmatrix} AA' (\Gamma_1 \Gamma_2) = \begin{bmatrix} \Gamma_1' AA' \Gamma_1 & \Gamma_1' AA' \Gamma_2 \\ \Gamma_2' AA' \Gamma_1 & \Gamma_2' AA' \Gamma_2 \end{bmatrix} = \begin{bmatrix} \Lambda_1^2 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\Gamma_1' AA' \Gamma_1 = \Lambda_1^2, \tag{5 A. 4·1}$$

$$\Gamma_2' AA' \Gamma_2 = 0. \tag{5 A. 4·2}$$

Pre- and post-multiply (5 A. 4·1) by Λ_1^{-1}:

$$\Lambda_1^{-1} \Gamma_1' AA' \Gamma_1 \Lambda_1^{-1} = I_r. \tag{5 A. 4·3}$$

Let

$$\Delta_1 = A' \Gamma_1 \Lambda_1^{-1} \quad (\Delta_1: q \times r)$$

and then

$$\Delta_1 \Lambda_1 = A' \Gamma_1, \tag{5 A. 4·4}$$

$$\Delta_1' \Delta_1 = I_r \quad \text{(by 5 A. 4·3)}. \tag{5 A. 4·5}$$

From (5 A. 4·2) $(\Gamma_2' A)(\Gamma_2' A)' = 0$ and hence $\Gamma_2' A = 0$.

From (5 A. 4·5) we see that the columns of Δ_1 are orthogonal and hence there exists a matrix Δ_2 such that

$$\Delta = [\Delta_1 \Delta_2] \tag{5 A. 4·6}$$

is orthogonal. The matrices Γ and Δ satisfy the theorem, since on performing the multiplication we find

$$\Gamma' A = (\Lambda 0) \Delta'. \quad \text{Q.E.D.}$$

5 A. 5. An elementary transformation on a matrix B is equivalent to $E_1 \dots E_n B E_{n+1} \dots E_m$ where some of the E's are of the type 5 A. 1, and the remainder are of type 5 A. 2. Hence if B is square and n.s. we can write (by 5 A. 4)

$$B = E_n^{-1} \dots E_1^{-1} E_m^{-1} \dots E_{n+1}^{-1}$$
$$\equiv F_m F_{m-1} \dots F_1,$$

where the F's are of type 5 A. 1 or 5 A. 2 (see Frazer *et al.* 1938, § 3·12).

5 A. 6. Diagonal matrices are commutative (see R. G. Frazer *et al.* 1938, § 1·6).

5 A. 7. If A and D are square, not necessarily of the same dimensions, and $|D| \neq 0$, then

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D|\,|A - BD^{-1}C| \quad \text{(see Hsu, 1941–2).}$$

Proof.
$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \left|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix}\right| = \left|\begin{bmatrix} A - BD^{-1}C & B \\ 0 & D \end{bmatrix}\right|$$
$$= |D|\,|A - BD^{-1}C|. \quad \text{Q.E.D.}$$

5 A. 8. $$|I + \bar{x}\bar{y}'| = 1 + \bar{y}'\bar{x} \quad \text{(see Aitken, 1939, § 37).}$$

5 A. 9. If $\bar{x}$ is $n \times 1$, there exists an orthogonal matrix Γ: $n \times n$ such that $\bar{x}'\Gamma = (0, \dots, 0, a)$, $a \geqslant 0$, with $a = 0$ if and only if $\bar{x}' = (0, \dots, 0)$.

Proof. If $\bar{x}' \neq (0, \dots, 0)$, choose Γ as any orthogonal matrix with elements in the last column

$$x_i \left(\sum_{i=1}^{n} x_i^2\right)^{-\frac{1}{2}} \quad (i = 1, \dots, n).$$

Then $\bar{x}'\Gamma = \left[0, \dots, 0, \left(\sum_1^n x_i^2\right)^{\frac{1}{2}}\right]$. If $\bar{x}' = (0, \dots, 0)$, any orthogonal matrix will do, and $a = 0$. Q.E.D.

5 A. 10. If M is any $m \times m$ matrix, there exists an orthogonal matrix Δ: $m \times m$, such that $M\Delta$ is triangular with 0's below the main diagonal, and with diagonal elements $\geqslant 0$.

Proof. By induction. If M is 1×1, then $\Delta = \pm 1$ will do. Assume true for a matrix of dimensions $(m-1) \times (m-1)$. Take Γ so that the product of the last row of M into Γ is $(0, \dots, 0, a)$, $a \geqslant 0$ (by 5 A. 9). Then

$$M\Gamma = \begin{bmatrix} M_1 & b \\ 0 & a \end{bmatrix} \quad (M_1: (m-1) \times (m-1);\ b: (m-1) \times 1;\ a: 1 \times 1).$$

By the hypothesis of the induction there exists an orthogonal matrix Γ_1 such that $M_1\Gamma_1$ is triangular with 0's below the main diagonal. Let

$$\Delta = \Gamma \begin{bmatrix} \Gamma_1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then
$$M\Delta = M\Gamma \begin{bmatrix} \Gamma_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} M_1 & b \\ 0 & a \end{bmatrix}\begin{bmatrix} \Gamma_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} M_1\Gamma_1 & b \\ 0 & a \end{bmatrix},$$

which is of the required form. Δ is the product of two orthogonal matrices and hence is orthogonal. Q.E.D.

5 A. 11. If D_1 is the adjoint of any determinant D, and M and M_1 are corresponding m-rowed minors of D and D_1 respectively, then M_1 is equal to the product of D^{m-1} by the algebraic complement of M (see Bôcher, 1938, § 11).

5 A. 12. If A: $m \times n$ and B: $n \times m$, then $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof. $\operatorname{tr}(AB) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}a_{ji} = \operatorname{tr}(BA)$. Q.E.D.

DEFINITION 5 A. 13*a*. If $f(x_1, \ldots, x_n)$ is a differentiable function of $x_1 \ldots x_n$, then

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$$

is called the differential of f.

DEFINITION 5 A. 13*b*. If $A = (a_{ij})$ $(i = 1, \ldots, p; j = 1, \ldots, q)$, where the a_{ij} are functions of $(x_1 \ldots x_r) = a_{ij}(x_1 \ldots x_r)$, then

$$dA = (da_{ij}) \quad (i = 1, \ldots, p; j = 1, \ldots, q).$$

5 A. 13. $$d(AB) = (dA)B + A(dB).$$

Proof. $d(AB) = d(\sum_j a_{ij}b_{jk}) = (\sum_j d(a_{ij}b_{jk}))$

$$= \sum_j [(da_{ij})b_{jk} + a_{ij}(db_{jk})] = (dA)B + A(dB). \quad \text{Q.E.D.}$$

5 A. 14. If A is square, n.s., $d|A| = |A| \operatorname{tr}[A^{-1}(dA)]$.

Proof. Let $A^{-1} = (a^{ij})$.

(5 A. 14·1) $$a^{ij} = (\text{cofactor } a_{ji})\,|A|^{-1},$$

(5 A. 14·2) $$d|A| = \sum_i \sum_j \frac{\partial |A|}{\partial a_{ij}} da_{ij} \quad \text{(by Definition 5 A. 13}a\text{)}.$$

(5 A. 14·3) $$\frac{\partial |A|}{\partial a_{ij}} = \text{cofactor } a_{ij} = |A| a^{ji} \quad \text{(by 5 A. 14·1)}.$$

Substitute (5 A. 14·3) in (5 A. 14·2):

(5 A. 14·4) $$d|A| = |A| \sum_i \sum_j a^{ji} da_{ij}.$$

The (i, i) term of $A^{-1}(dA)$ is $\sum_j a^{ij}(da_{ji})$, and $\operatorname{tr} A^{-1}(dA) = \sum_i \sum_j a^{ij}(da_{ji})$, which when put in (5 A. 14·4) gives the result of the theorem. Q.E.D.

5 A. 15. If A is square, n.s., $(dA^{-1}) = -A^{-1}(dA)A^{-1}$.

Proof. $d(AA^{-1}) = (dA)A^{-1} + A(dA^{-1}) = 0$. Q.E.D.

5 A. 16. A p.d. matrix $\overline{V}$: $p \times p$ can be factored uniquely into $\overline{V} = \overline{T}\overline{T}'$, $t_{ii} > 0$ $(i = 1, \ldots, p)$ (see Mahalanobis, Bose & Roy, 1937).

5 A. 17. If U: $p \times p$ is n.s., then UU' is p.d. (see Birkhoff & MacLane, 1947, p. 246).

5 A. 18. If Γ is orthogonal and $|\Gamma + I| \neq 0$, there is a one-to-one relation between Γ and $\hat{X}$:

$$\Gamma = 2(\hat{X} + I)^{-1} - I,$$

$$\hat{X} = 2(\Gamma + I)^{-1} - I.$$

Proof. Show that $\hat{X}$ is skew symmetric, i.e. $\hat{X} + \hat{X}' = 0$.

5 A. 19. UU' of 5 A. 17 can be expressed uniquely as $\tilde{T}\tilde{T}'$, where $t_{ii}>0$ $(i=1,\ldots,p)$. If $|\tilde{T}^{-1}U+I|\neq 0$, U can be expressed uniquely in the form

$$U=\tilde{T}[2(\hat{X}+I)^{-1}-I].$$

Proof. Follows from 5 A. 16, 5 A. 17 and 5 A. 18.

5 A. 20. Given Γ as in 5 A. 18, then

$$d\Gamma=-\tfrac{1}{2}(\Gamma+I)(d\hat{X})(\Gamma+I).$$

5 A. 21. Let

$$\overline{V}=\begin{bmatrix}\overline{V}_{11} & V_{12}\\ V_{21} & \overline{V}_{22}\end{bmatrix}$$

be p.d. ($\overline{V}_{11}$: $p\times p$; V_{12}: $p\times q$; $\overline{V}_{22}$: $q\times q$) ($p\leqslant q$). Then there exists a matrix G of the form

$$\begin{bmatrix}G_1 & 0\\ 0 & G_2\end{bmatrix}\quad (G_1\colon p\times p;\ G_2\colon q\times q)$$

such that

$$G\overline{V}G'=\Phi=\begin{bmatrix}I_p & D_r & 0\\ D_r & & \\ 0 & & I_q\end{bmatrix},$$

where D_r: $p\times p$ has elements $r_1,\ldots,r_p$, and $r_1^2\ldots r_p^2$ are the roots of the determinantal equation $|V_{12}\overline{V}_{22}^{-1}V_{21}-\theta\overline{V}_{11}|=0$ (see H. Hotelling, 1936).

Proof. Since $\overline{V}$ is p.d., $\overline{V}_{11}$ and $\overline{V}_{22}$ are of rank p and q respectively, and hence by 5 A. 4 there exist two non-singular matrices F_1 and F_2 such that

$$F_1\overline{V}_{11}F_1'=I_p, \tag{5 A. 21·1}$$

$$F_2\overline{V}_{22}F_2'=I_q. \tag{5 A. 21·2}$$

Let $A=F_1V_{12}F_2'$; then by 5 A. 4 there exist two orthogonal matrices Γ_1 and Γ_2 such that

$$\Gamma_1A\Gamma_2'=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_p\end{bmatrix}, \tag{5 A. 21·3}$$

where $\lambda_1^2,\ldots,\lambda_p^2$ are the characteristic roots of AA', i.e. roots of $|AA'-\lambda I|=0$.

Pre- and post-multiply (5 A. 21·1) by Γ_1 and Γ_1' respectively, and pre- and post-multiply (5 A. 21·2) by Γ_2 and Γ_2' respectively. Then

$$\Gamma_1F_1\overline{V}_{11}F_1'\Gamma_1'=\Gamma_1I_p\Gamma_1'=I_p, \tag{5 A. 21·4}$$

$$\Gamma_2F_2\overline{V}_{22}F_2'\Gamma_2'=\Gamma_2I_q\Gamma_2'=I_q. \tag{5 A. 21·5}$$

But $AA'=F_1V_{12}F_2'F_2V_{21}F_1'$, and from (5 A. 21·1) and (5 A. 21·2)

$$\overline{V}_{11}=F_1^{-1}F_1'^{-1}\quad\text{and}\quad\overline{V}_{22}^{-1}=F_2'F_2.$$

Hence

$$\begin{aligned}AA'-\lambda I&=F_1[V_{12}\overline{V}_{22}^{-1}V_{21}-\lambda F_1^{-1}F_1'^{-1}]F_1'\\ &=F_1[V_{12}\overline{V}_{22}^{-1}V_{21}-\lambda\overline{V}_{11}]F_1'.\end{aligned} \tag{5 A. 21·6}$$

F_1 is n.s., so $|AA'-\lambda I|=0$ is equivalent to $|V_{12}\overline{V}_{22}^{-1}V_{21}-\lambda\overline{V}_{11}|=0$, and hence $\lambda_1^2,\ldots,\lambda_p^2$ are the roots of this last equation. Hence

$$\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_p\end{bmatrix}=(D_r\,0). \tag{5 A. 21·7}$$

Let $G_1 = \Gamma_1 F_1$ and $G_2 = \Gamma_2 F_2$; then G_1 and G_2 are the matrices of the theorem, since by substituting these values in G and multiplying out $G\overline{V}G'$ we get

$$G\overline{V}G' = \begin{bmatrix} \Gamma_1 F_1 \overline{V}_{11} F_1' \Gamma_1' & \Gamma_1 F_1 V_{12} F_2' \Gamma_2' \\ \Gamma_2 F_2 V_{21} F_1' \Gamma_1' & \Gamma_2 F_2 \overline{V}_{22} F_2' \Gamma_2' \end{bmatrix}, \tag{5 A. 21·8}$$

which by (5 A. 21·4), (5 A. 21·5) and (5 A. 21·7) is

$$\begin{bmatrix} I_p & D_r & 0 \\ D_r & & \\ 0 & & I_q \end{bmatrix}. \quad \text{Q.E.D.}$$

B. *General properties of Jacobians*

5 B. 1. $$D(Y; X) = \frac{1}{D(X; Y)}, \text{ if } D(X; Y) \neq 0.$$

5 B. 2. $D(X; Y) = D(X; U) D(U; V) D(V; Y)$, where U and V are any functions of X and Y such that none of the terms on the right vanish. The extension of this to more variables and to the introduction of functions of U and V is immediate.

5 B. 3. Given a matrix transformation, linear or not, $Y = F(X)$. Then the transformation of the differentials, $dY = dF(X)$, is linear in the differentials, and the J of the original transformation is the J of the linear transformation in the differentials.

Proof. $Y = F(X)$ implies a set of transformations in the elements of Y:

$$y_i = F_i(x_1 \dots x_m) \quad (i = 1, \dots, m).$$

Let $$(dy)' = (dy_1 \dots dy_m); \quad (dx)' = (dx_1 \dots dx_m).$$

Then the equation $$dY = dF(X)$$

is of the form $$(dy) = G(dx),$$

where $$G = \left[\frac{\partial f_i}{\partial x_j}\right] \quad (i, j = 1, \dots, m),$$

and G is by definition the J of the original transformation. Q.E.D.

5 B. 4. If $X = F_1(U)$ and $Y = F_2(V)$ are a transformation from variables X and Y to new variables U and V, then

$$J = D(U, V; X, Y) = D(U; X) D(V; Y).$$

We wish to express our gratitude to Prof. Harold Hotelling for his help and encouragement in the preparation of this paper.

We are grateful to Dr Ralph A. Bradley for the use of his lecture notes.

REFERENCES

Aitken, A. C. (1939). *Determinants and Matrices.* Edinburgh: Oliver and Boyd.
Albert, A. A. (1946). *Introduction to Algebraic Theories.* Chicago: University of Chicago Press.
Birkhoff, G. & MacLane, S. (1947). *A Survey of Modern Algebra.* New York: Macmillan Co.
Bôcher, M. (1938). *Introduction to Higher Algebra.* New York: Macmillan Co.
Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton: Princeton University Press.

FISHER, R. A. (1939). The sampling distribution of some statistics obtained from non-linear equations. *Ann. Eugen., Lond.*, **9**, 238.

FRAZER, R. G., DUNCAN, W. J. & COLLAR, A. R. (1938). *Elementary Matrices and some Applications to Dynamics and Differential Equations.* Cambridge University Press.

HOTELLING, H. (1936). Relations between two sets of variates. *Biometrika*, **28**, 321.

HSU, P. L. (1939). On the distribution of roots of certain determinantal equations. *Ann. Eugen., Lond.*, **9**, 250.

HSU, P. L. (1941–2). On the limiting distribution of the canonical correlations. *Biometrika*, **32**, 38.

MAHALANOBIS, P. C., BOSE, R. C. & ROY, S. N. (1937). Normalization of statistical variates and the use of rectangular coordinates in the theory of sampling distributions. *Sankhyā*, **3**, 1.

ROY, S. N. (1938–40). p-Statistics, or some generalizations in analysis of variance appropriate to multivariate problems. *Sankhyā*, **4**, 381.

TWO APPLICATIONS OF BIVARIATE k-STATISTICS

By M. B. COOK

University College, London

1. In a previous paper (Cook, 1951) the derivation of bivariate k-statistics and product cumulant formulae up to weight 6 was discussed. In the present paper these formulae are used to find approximations to the moments of the sampling distributions of the correlation coefficient and the regression coefficients under the very general assumption that the bivariate surface generating the samples has as many cumulants as desired.

2. *The sample correlation coefficient.* It is assumed that there are two correlated variables x and y. A sample of n pairs of observations $(x_i y_i)$ is available, whence the sample correlation coefficient is commonly written

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{[\Sigma(x-\bar{x})^2\Sigma(y-\bar{y})^2]^{\frac{1}{2}}},$$

the summation in each case extending over n. This expression is, in terms of bivariate k-statistics and population cumulants,

$$r = \frac{k_{11}}{(k_{02}k_{20})^{\frac{1}{2}}} = \frac{\kappa_{11}}{(\kappa_{20}\kappa_{02})^{\frac{1}{2}}}\left(1+\frac{k_{11}-\kappa_{11}}{\kappa_{11}}\right)\left(1+\frac{k_{20}-\kappa_{20}}{\kappa_{20}}\right)^{-\frac{1}{2}}\left(1+\frac{k_{02}-\kappa_{02}}{\kappa_{02}}\right)^{-\frac{1}{2}}. \tag{1}$$

This expression can be expanded as the product of two binomials, a process which, it can be argued, will be valid in all but a very small proportion of cases provided n be large. Expanding and taking expectations we have

$$\begin{aligned}
\mathscr{E}(r) = \rho\Big\{&1-\frac{1}{2}\frac{1}{\kappa_{11}\kappa_{20}}\kappa\begin{pmatrix}1&2\\1&0\end{pmatrix}-\frac{1}{2}\frac{1}{\kappa_{11}\kappa_{02}}\kappa\begin{pmatrix}1&0\\1&2\end{pmatrix}+\frac{1}{4}\frac{1}{\kappa_{20}\kappa_{02}}\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}+\frac{3}{8}\frac{1}{\kappa_{20}^2}\kappa\begin{pmatrix}2&2\\0&0\end{pmatrix}\\
&+\frac{3}{8}\frac{1}{\kappa_{02}^2}\kappa\begin{pmatrix}0&0\\2&2\end{pmatrix}+\frac{3}{8}\frac{1}{\kappa_{11}\kappa_{20}^2}\kappa\begin{pmatrix}1&2&2\\1&0&0\end{pmatrix}+\frac{3}{8}\frac{1}{\kappa_{11}\kappa_{02}^2}\kappa\begin{pmatrix}1&0&0\\1&2&2\end{pmatrix}+\frac{1}{4}\frac{1}{\kappa_{11}\kappa_{20}\kappa_{02}}\kappa\begin{pmatrix}1&2&0\\1&0&2\end{pmatrix}\\
&-\frac{3}{16}\frac{1}{\kappa_{20}^2\kappa_{02}}\kappa\begin{pmatrix}2&2&0\\0&0&2\end{pmatrix}-\frac{3}{16}\frac{1}{\kappa_{20}\kappa_{02}^2}\kappa\begin{pmatrix}2&0&0\\0&2&2\end{pmatrix}-\frac{5}{16}\frac{1}{\kappa_{20}^3}\kappa\begin{pmatrix}2&2&2\\0&0&0\end{pmatrix}-\frac{5}{16}\frac{1}{\kappa_{02}^3}\kappa\begin{pmatrix}0&0&0\\2&2&2\end{pmatrix}\\
&-\frac{15}{16}\frac{1}{\kappa_{11}\kappa_{20}^3}\kappa\begin{pmatrix}1&2\\1&0\end{pmatrix}\kappa\begin{pmatrix}2&2\\0&0\end{pmatrix}-\frac{15}{16}\frac{1}{\kappa_{11}\kappa_{02}^3}\kappa\begin{pmatrix}1&0\\1&2\end{pmatrix}\kappa\begin{pmatrix}0&0\\2&2\end{pmatrix}-\frac{3}{8}\frac{1}{\kappa_{11}\kappa_{20}^2\kappa_{02}}\kappa\begin{pmatrix}1&2\\1&0\end{pmatrix}\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}\\
&-\frac{3}{16}\frac{1}{\kappa_{11}\kappa_{20}^2\kappa_{02}}\kappa\begin{pmatrix}2&2\\0&0\end{pmatrix}\kappa\begin{pmatrix}1&0\\1&2\end{pmatrix}-\frac{3}{8}\frac{1}{\kappa_{11}\kappa_{20}\kappa_{02}^2}\kappa\begin{pmatrix}1&0\\1&2\end{pmatrix}\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}-\frac{3}{16}\frac{1}{\kappa_{11}\kappa_{20}\kappa_{02}^2}\kappa\begin{pmatrix}0&0\\2&2\end{pmatrix}\kappa\begin{pmatrix}1&2\\1&0\end{pmatrix}\\
&+\frac{15}{32}\frac{1}{\kappa_{20}^3\kappa_{02}}\kappa\begin{pmatrix}2&2\\0&0\end{pmatrix}\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}+\frac{15}{32}\frac{1}{\kappa_{20}\kappa_{02}^3}\kappa\begin{pmatrix}2&0\\0&2\end{pmatrix}\kappa\begin{pmatrix}0&0\\2&2\end{pmatrix}+\frac{9}{64}\frac{1}{\kappa_{20}^2\kappa_{02}^2}\kappa\begin{pmatrix}2&2\\0&0\end{pmatrix}\kappa\begin{pmatrix}0&0\\2&2\end{pmatrix}\\
&+\frac{9}{32}\frac{1}{\kappa_{20}^2\kappa_{02}^2}\kappa^2\begin{pmatrix}2&0\\0&2\end{pmatrix}+\frac{105}{128}\frac{1}{\kappa_{20}^4}\kappa^2\begin{pmatrix}2&2\\0&0\end{pmatrix}+\frac{105}{128}\frac{1}{\kappa_{02}^4}\kappa^2\begin{pmatrix}0&0\\2&2\end{pmatrix}\Big\},
\end{aligned} \tag{2}$$

where ρ is the parent correlation. Substituting for the cumulants we have on reduction

$$\mu_1'(r) = \rho\left\{1 - \frac{1}{2n} - \frac{3}{8n^2} + \left(\frac{3}{8n} - \frac{15}{16n^2}\right)\frac{\kappa_{40}}{\kappa_{20}^2} + \left(\frac{3}{8n} - \frac{15}{16n^2}\right)\frac{\kappa_{04}}{\kappa_{02}^2} - \left(\frac{1}{2n} - \frac{5}{4n^2}\right)\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}}\right.$$
$$- \left(\frac{1}{2n} - \frac{5}{4n^2}\right)\frac{\kappa_{13}}{\kappa_{11}\kappa_{02}} + \left(\frac{1}{4n} + \frac{7}{8n^2}\right)\frac{\kappa_{22}}{\kappa_{20}\kappa_{02}} + \left(\frac{1}{2n} - \frac{3}{4n^2}\right)\frac{\kappa_{11}^2}{\kappa_{20}\kappa_{02}} - \frac{5}{16n^2}\frac{\kappa_{60}}{\kappa_{20}^3} - \frac{5}{16n^2}\frac{\kappa_{06}}{\kappa_{02}^3}$$
$$- \frac{5}{4n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3} - \frac{5}{4n^2}\frac{\kappa_{03}^2}{\kappa_{02}^3} - \frac{3}{16n^2}\frac{\kappa_{42}}{\kappa_{20}^2\kappa_{02}} - \frac{3}{16n^2}\frac{\kappa_{24}}{\kappa_{20}\kappa_{02}^2} - \frac{9}{4n^2}\frac{\kappa_{11}\kappa_{31}}{\kappa_{20}^2\kappa_{02}} - \frac{9}{4n^2}\frac{\kappa_{11}\kappa_{13}}{\kappa_{20}\kappa_{02}^2} - \frac{3}{4n^2}\frac{\kappa_{21}^2}{\kappa_{20}^2\kappa_{02}}$$
$$- \frac{3}{4n^2}\frac{\kappa_{12}^2}{\kappa_{20}\kappa_{02}^2} + \frac{3}{8n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2} + \frac{3}{8n^2}\frac{\kappa_{15}}{\kappa_{11}\kappa_{02}^2} + \frac{3}{2n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2} + \frac{3}{2n^2}\frac{\kappa_{03}\kappa_{12}}{\kappa_{11}\kappa_{02}^2} + \frac{1}{4n^2}\frac{\kappa_{33}}{\kappa_{11}\kappa_{20}\kappa_{02}}$$
$$+ \frac{1}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}\kappa_{20}\kappa_{02}} + \frac{105}{128n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4} + \frac{105}{128n^2}\frac{\kappa_{04}^2}{\kappa_{02}^4} + \frac{15}{32n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{20}^3\kappa_{02}} + \frac{15}{32n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{20}\kappa_{02}^3} + \frac{15}{16n^2}\frac{\kappa_{40}\kappa_{11}^2}{\kappa_{20}^3\kappa_{02}}$$
$$+ \frac{15}{16n^2}\frac{\kappa_{04}\kappa_{11}^2}{\kappa_{20}\kappa_{02}^3} + \frac{9}{32n^2}\frac{\kappa_{22}^2}{\kappa_{02}^2\kappa_{20}^2} + \frac{9}{8n^2}\frac{\kappa_{11}^4}{\kappa_{20}^2\kappa_{02}^2} + \frac{9}{8n^2}\frac{\kappa_{22}\kappa_{11}^2}{\kappa_{20}^2\kappa_{02}^2} + \frac{9}{64n^2}\frac{\kappa_{40}\kappa_{04}}{\kappa_{20}^2\kappa_{02}^2} - \frac{15}{16n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}$$
$$\left. - \frac{15}{16n^2}\frac{\kappa_{04}\kappa_{13}}{\kappa_{11}\kappa_{02}^3} - \frac{3}{8n^2}\frac{\kappa_{22}\kappa_{13}}{\kappa_{11}\kappa_{20}\kappa_{02}^2} - \frac{3}{8n^2}\frac{\kappa_{22}\kappa_{31}}{\kappa_{11}\kappa_{20}^2\kappa_{02}} - \frac{3}{16n^2}\frac{\kappa_{31}\kappa_{04}}{\kappa_{11}\kappa_{20}^2\kappa_{02}} - \frac{3}{16n^2}\frac{\kappa_{15}\kappa_{40}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}\right\}. \qquad (3)$$

Again by expanding

$$\rho^2\left(1 + \frac{k_{11} - \kappa_{11}}{\kappa_{11}}\right)^2\left(1 + \frac{k_{20} - \kappa_{20}}{\kappa_{20}}\right)^{-1}\left(1 + \frac{k_{02} - \kappa_{02}}{\kappa_{02}}\right)^{-1}$$

and taking expectations, we have, to order n^{-2},

$$\mu_2'(r) = \rho^2\left\{1 - \frac{3}{n} + \frac{3}{n^2} + \left(\frac{1}{n} - \frac{5}{n^2}\right)\frac{\kappa_{40}}{\kappa_{20}^2} + \left(\frac{1}{n} - \frac{5}{n^2}\right)\frac{\kappa_{04}}{\kappa_{02}^2} + \left(\frac{1}{n} + \frac{7}{n^2}\right)\frac{\kappa_{22}}{\kappa_{20}\kappa_{02}} + \left(\frac{2}{n} - \frac{12}{n^2}\right)\frac{\kappa_{11}^2}{\kappa_{20}\kappa_{02}}\right.$$
$$+ \left(\frac{10}{n^2} - \frac{2}{n}\right)\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}} + \left(\frac{10}{n^2} - \frac{2}{n}\right)\frac{\kappa_{13}}{\kappa_{11}\kappa_{02}} + \left(\frac{1}{n} - \frac{5}{n^2}\right)\frac{\kappa_{22}}{\kappa_{11}^2} + \left(\frac{1}{n} + \frac{1}{n^2}\right)\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2} - \frac{1}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3} - \frac{1}{n^2}\frac{\kappa_{06}}{\kappa_{02}^3}$$
$$- \frac{4}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3} - \frac{4}{n^2}\frac{\kappa_{03}^2}{\kappa_{02}^3} - \frac{1}{n^2}\frac{\kappa_{42}}{\kappa_{20}^2\kappa_{02}} - \frac{1}{n^2}\frac{\kappa_{24}}{\kappa_{20}\kappa_{02}^2} - \frac{16}{n^2}\frac{\kappa_{11}\kappa_{31}}{\kappa_{20}^2\kappa_{02}} - \frac{16}{n^2}\frac{\kappa_{11}\kappa_{13}}{\kappa_{20}\kappa_{02}^2} - \frac{4}{n^2}\frac{\kappa_{21}^2}{\kappa_{20}^2\kappa_{02}}$$
$$- \frac{4}{n^2}\frac{\kappa_{12}^2}{\kappa_{20}\kappa_{02}^2} + \frac{2}{n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2} + \frac{2}{n^2}\frac{\kappa_{15}}{\kappa_{11}\kappa_{02}^2} + \frac{8}{n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2} + \frac{8}{n^2}\frac{\kappa_{03}\kappa_{12}}{\kappa_{11}\kappa_{02}^2} + \frac{2}{n^2}\frac{\kappa_{33}}{\kappa_{11}\kappa_{20}\kappa_{02}} + \frac{8}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}\kappa_{20}\kappa_{02}}$$
$$- \frac{1}{n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}} - \frac{1}{n^2}\frac{\kappa_{24}}{\kappa_{11}^2\kappa_{02}} - \frac{2}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}} - \frac{2}{n^2}\frac{\kappa_{03}\kappa_{21}}{\kappa_{11}^2\kappa_{02}} - \frac{2}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}} - \frac{2}{n^2}\frac{\kappa_{12}^2}{\kappa_{11}^2\kappa_{02}} + \frac{3}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}$$
$$+ \frac{3}{n^2}\frac{\kappa_{04}^2}{\kappa_{02}^4} + \frac{3}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{20}^3\kappa_{02}} + \frac{3}{n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{20}\kappa_{02}^3} + \frac{6}{n^2}\frac{\kappa_{40}\kappa_{11}^2}{\kappa_{20}^3\kappa_{02}} + \frac{6}{n^2}\frac{\kappa_{04}\kappa_{11}^2}{\kappa_{20}\kappa_{02}^3} + \frac{2}{n^2}\frac{\kappa_{22}^2}{\kappa_{20}^2\kappa_{02}^2}$$
$$+ \frac{8}{n^2}\frac{\kappa_{22}\kappa_{11}^2}{\kappa_{20}^2\kappa_{02}^2} + \frac{8}{n^2}\frac{\kappa_{11}^4}{\kappa_{20}^2\kappa_{02}^2} + \frac{1}{n^2}\frac{\kappa_{40}\kappa_{04}}{\kappa_{20}^2\kappa_{02}^2} - \frac{6}{n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3} - \frac{6}{n^2}\frac{\kappa_{04}\kappa_{13}}{\kappa_{11}\kappa_{02}^3} - \frac{4}{n^2}\frac{\kappa_{22}\kappa_{31}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}$$
$$- \frac{4}{n^2}\frac{\kappa_{22}\kappa_{13}}{\kappa_{11}\kappa_{20}\kappa_{02}^2} - \frac{2}{n^2}\frac{\kappa_{40}\kappa_{13}}{\kappa_{11}\kappa_{20}^2\kappa_{02}} - \frac{2}{n^2}\frac{\kappa_{04}\kappa_{31}}{\kappa_{11}\kappa_{20}\kappa_{02}^2} + \frac{2}{n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}} + \frac{2}{n^2}\frac{\kappa_{13}^2}{\kappa_{11}^2\kappa_{02}} + \frac{1}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}}$$
$$\left. + \frac{1}{n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{11}^2\kappa_{02}} + \frac{2}{n^2}\frac{\kappa_{31}\kappa_{13}}{\kappa_{11}^2\kappa_{20}\kappa_{02}} + \frac{1}{n^2}\frac{\kappa_{22}^2}{\kappa_{11}^2\kappa_{20}\kappa_{02}}\right\}. \qquad (4)$$

The crude third and fourth moments to order n^{-2} follow in similar fashion:

$$\begin{aligned}\mu'_3(r) = \rho^3\Big\{&1-\frac{15}{2n}+\frac{261}{8n^2}+\left(\frac{15}{8n}-\frac{285}{16n^2}\right)\frac{\kappa_{40}}{\kappa_{20}^2}+\left(\frac{15}{8n}-\frac{285}{16n^2}\right)\frac{\kappa_{04}}{\kappa_{02}^2}+\left(\frac{9}{4n}+\frac{189}{8n^2}\right)\frac{\kappa_{22}}{\kappa_{20}\kappa_{02}}\\
&+\left(\frac{9}{2n}-\frac{225}{4n^2}\right)\frac{\kappa_{11}^2}{\kappa_{20}\kappa_{02}}+\left(\frac{171}{4n^2}-\frac{9}{2n}\right)\frac{\kappa_{31}}{\kappa_{20}\kappa_{11}}+\left(\frac{171}{4n^2}-\frac{9}{2n}\right)\frac{\kappa_{13}}{\kappa_{02}\kappa_{11}}+\left(\frac{3}{n}-\frac{111}{4n^2}\right)\frac{\kappa_{22}}{\kappa_{11}^2}\\
&+\left(\frac{3}{n}-\frac{9}{2n^2}\right)\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}-\frac{35}{16n^2}\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{35}{16n^2}\frac{\kappa_{06}}{\kappa_{02}^3}-\frac{35}{4n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}-\frac{35}{4n^2}\frac{\kappa_{03}^2}{\kappa_{02}^3}-\frac{45}{16n^2}\frac{\kappa_{42}}{\kappa_{20}^2\kappa_{02}}-\frac{45}{16n^2}\frac{\kappa_{24}}{\kappa_{20}\kappa_{02}^2}\\
&-\frac{225}{4n^2}\frac{\kappa_{11}\kappa_{31}}{\kappa_{20}^2\kappa_{02}}-\frac{225}{4n^2}\frac{\kappa_{11}\kappa_{13}}{\kappa_{20}\kappa_{02}^2}-\frac{45}{4n^2}\frac{\kappa_{21}^2}{\kappa_{20}^2\kappa_{02}}-\frac{45}{4n^2}\frac{\kappa_{12}^2}{\kappa_{20}\kappa_{02}^2}+\frac{45}{8n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}+\frac{45}{8n^2}\frac{\kappa_{15}}{\kappa_{11}\kappa_{02}^2}+\frac{45}{2n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}\\
&+\frac{45}{2n^2}\frac{\kappa_{03}\kappa_{12}}{\kappa_{11}\kappa_{02}^2}+\frac{27}{4n^2}\frac{\kappa_{33}}{\kappa_{11}\kappa_{20}\kappa_{02}}+\frac{27}{n^2}\frac{\kappa_{21}\kappa_{12}}{\kappa_{11}\kappa_{20}\kappa_{02}}-\frac{9}{2n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}-\frac{9}{2n^2}\frac{\kappa_{24}}{\kappa_{11}^2\kappa_{02}}+\frac{9}{8n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}\\
&+\frac{9}{8n^2}\frac{\kappa_{04}\kappa_{20}}{\kappa_{11}^2\kappa_{02}}-\frac{9}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}}-\frac{9}{n^2}\frac{\kappa_{03}\kappa_{21}}{\kappa_{11}^2\kappa_{02}}-\frac{9}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}}-\frac{9}{n^2}\frac{\kappa_{12}^2}{\kappa_{11}^2\kappa_{02}}+\frac{1}{n^2}\frac{\kappa_{33}}{\kappa_{11}^3}\\
&-\frac{3}{2n^2}\frac{\kappa_{20}\kappa_{13}}{\kappa_{11}^3}-\frac{3}{2n^2}\frac{\kappa_{02}\kappa_{31}}{\kappa_{11}^3}+\frac{1}{n^2}\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}^3}+\frac{3}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}^3}+\frac{945}{128n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}+\frac{945}{128n^2}\frac{\kappa_{04}^2}{\kappa_{02}^4}+\frac{315}{32n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{20}^3\kappa_{02}}\\
&+\frac{315}{32n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{20}\kappa_{02}^3}+\frac{315}{16n^2}\frac{\kappa_{40}\kappa_{11}^2}{\kappa_{20}^3\kappa_{02}}+\frac{315}{16n^2}\frac{\kappa_{04}\kappa_{11}^2}{\kappa_{20}\kappa_{02}^3}+\frac{225}{32n^2}\frac{\kappa_{22}^2}{\kappa_{20}^2\kappa_{02}^2}+\frac{225}{8n^2}\frac{\kappa_{22}\kappa_{11}^2}{\kappa_{20}^2\kappa_{02}^2}+\frac{225}{8n^2}\frac{\kappa_{11}^4}{\kappa_{20}^2\kappa_{02}^2}\\
&+\frac{225}{64n^2}\frac{\kappa_{40}\kappa_{04}}{\kappa_{20}^2\kappa_{02}^2}-\frac{315}{16n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}-\frac{315}{16n^2}\frac{\kappa_{04}\kappa_{13}}{\kappa_{11}\kappa_{02}^3}-\frac{135}{8n^2}\frac{\kappa_{22}\kappa_{31}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}-\frac{135}{8n^2}\frac{\kappa_{22}\kappa_{13}}{\kappa_{11}\kappa_{20}\kappa_{02}^2}\\
&+\frac{135}{16n^2}\frac{\kappa_{40}\kappa_{13}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}-\frac{135}{16n^2}\frac{\kappa_{04}\kappa_{31}}{\kappa_{11}\kappa_{20}\kappa_{02}^2}+\frac{45}{4n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}^2}+\frac{45}{4n^2}\frac{\kappa_{13}^2}{\kappa_{11}^2\kappa_{02}^2}+\frac{45}{8n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}\\
&+\frac{45}{8n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{11}^2\kappa_{02}^2}+\frac{27}{2n^2}\frac{\kappa_{31}\kappa_{13}}{\kappa_{11}^2\kappa_{20}\kappa_{02}}-\frac{9}{2n^2}\frac{\kappa_{31}\kappa_{22}}{\kappa_{11}^3\kappa_{20}}-\frac{9}{2n^2}\frac{\kappa_{13}\kappa_{22}}{\kappa_{11}^3\kappa_{02}}+\frac{27}{4n^2}\frac{\kappa_{22}^2}{\kappa_{11}^2\kappa_{20}\kappa_{02}}\Big\},\end{aligned}\tag{5}$$

$$\begin{aligned}\mu'_4(r) = \rho^4\Big\{&1-\frac{14}{n}+\frac{129}{n^2}+\left(\frac{3}{n}-\frac{48}{n^2}\right)\frac{\kappa_{40}}{\kappa_{20}^2}+\left(\frac{3}{n}-\frac{48}{n^2}\right)\frac{\kappa_{04}}{\kappa_{02}^2}+\left(\frac{4}{n}+\frac{56}{n^2}\right)\frac{\kappa_{22}}{\kappa_{20}\kappa_{02}}+\left(\frac{8}{n}-\frac{168}{n^2}\right)\frac{\kappa_{11}^2}{\kappa_{20}\kappa_{02}}\\
&+\left(\frac{128}{n^2}-\frac{8}{n}\right)\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}}+\left(\frac{128}{n^2}-\frac{8}{n}\right)\frac{\kappa_{13}}{\kappa_{11}\kappa_{02}}+\left(\frac{6}{n}-\frac{90}{n^2}\right)\frac{\kappa_{22}}{\kappa_{11}^2}+\left(\frac{6}{n}-\frac{36}{n^2}\right)\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}-\frac{4}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{4}{n^2}\frac{\kappa_{06}}{\kappa_{02}^3}\\
&-\frac{16}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}-\frac{16}{n^2}\frac{\kappa_{03}^2}{\kappa_{02}^3}-\frac{6}{n^2}\frac{\kappa_{42}}{\kappa_{20}^2\kappa_{02}}-\frac{6}{n^2}\frac{\kappa_{24}}{\kappa_{20}\kappa_{02}^2}-\frac{144}{n^2}\frac{\kappa_{11}\kappa_{31}}{\kappa_{20}^2\kappa_{02}}-\frac{144}{n^2}\frac{\kappa_{11}\kappa_{13}}{\kappa_{20}\kappa_{02}^2}-\frac{24}{n^2}\frac{\kappa_{21}^2}{\kappa_{20}^2\kappa_{02}}\\
&-\frac{24}{n^2}\frac{\kappa_{12}^2}{\kappa_{20}\kappa_{02}^2}+\frac{12}{n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}+\frac{12}{n^2}\frac{\kappa_{15}}{\kappa_{11}\kappa_{02}^2}+\frac{48}{n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}+\frac{48}{n^2}\frac{\kappa_{03}\kappa_{12}}{\kappa_{11}\kappa_{02}^2}+\frac{16}{n^2}\frac{\kappa_{33}}{\kappa_{11}\kappa_{20}\kappa_{02}}+\frac{64}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}\kappa_{20}\kappa_{02}}\\
&-\frac{12}{n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}-\frac{12}{n^2}\frac{\kappa_{24}}{\kappa_{11}^2\kappa_{02}}+\frac{6}{n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}+\frac{6}{n^2}\frac{\kappa_{04}\kappa_{20}}{\kappa_{11}^2\kappa_{02}}-\frac{24}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}}-\frac{24}{n^2}\frac{\kappa_{03}\kappa_{21}}{\kappa_{11}^2\kappa_{02}}-\frac{24}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}}\\
&-\frac{24}{n^2}\frac{\kappa_{12}^2}{\kappa_{11}^2\kappa_{02}}+\frac{4}{n^2}\frac{\kappa_{33}}{\kappa_{11}^3}-\frac{12}{n^2}\frac{\kappa_{20}\kappa_{13}}{\kappa_{11}^3}-\frac{12}{n^2}\frac{\kappa_{02}\kappa_{31}}{\kappa_{11}^3}+\frac{4}{n^2}\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}^3}+\frac{12}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}^3}+\frac{15}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}\end{aligned}$$

$$+\frac{15\kappa_{04}^2}{n^2\kappa_{02}^4}+\frac{24}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{20}^3\kappa_{02}}+\frac{24}{n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{20}\kappa_{02}^3}+\frac{48}{n^2}\frac{\kappa_{40}\kappa_{11}^2}{\kappa_{20}^3\kappa_{02}}+\frac{48}{n^2}\frac{\kappa_{04}\kappa_{11}^2}{\kappa_{20}\kappa_{02}^3}+\frac{18}{n^2}\frac{\kappa_{22}^2}{\kappa_{20}^2\kappa_{02}^2}$$

$$+\frac{72}{n^2}\frac{\kappa_{22}\kappa_{11}^2}{\kappa_{20}^2\kappa_{02}^2}+\frac{72}{n^2}\frac{\kappa_{11}^4}{\kappa_{20}^2\kappa_{02}^2}+\frac{9}{n^2}\frac{\kappa_{40}\kappa_{04}}{\kappa_{20}^2\kappa_{02}^2}-\frac{48}{n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}-\frac{48}{n^2}\frac{\kappa_{04}\kappa_{13}}{\kappa_{11}\kappa_{02}^3}-\frac{48}{n^2}\frac{\kappa_{31}\kappa_{22}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}$$

$$-\frac{48}{n^2}\frac{\kappa_{13}\kappa_{22}}{\kappa_{11}\kappa_{20}\kappa_{02}^2}-\frac{24}{n^2}\frac{\kappa_{40}\kappa_{13}}{\kappa_{11}\kappa_{20}^2\kappa_{02}}-\frac{24}{n^2}\frac{\kappa_{04}\kappa_{31}}{\kappa_{11}\kappa_{20}\kappa_{02}^2}+\frac{36}{n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}^2}+\frac{36}{n^2}\frac{\kappa_{13}^2}{\kappa_{11}^2\kappa_{02}^2}+\frac{18}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}$$

$$+\frac{18}{n^2}\frac{\kappa_{04}\kappa_{22}}{\kappa_{11}^2\kappa_{02}^2}+\frac{48}{n^2}\frac{\kappa_{31}\kappa_{13}}{\kappa_{11}^2\kappa_{20}\kappa_{02}}+\frac{24}{n^2}\frac{\kappa_{22}^2}{\kappa_{11}^2\kappa_{20}\kappa_{02}}-\frac{24}{n^2}\frac{\kappa_{31}\kappa_{22}}{\kappa_{11}^3\kappa_{20}}-\frac{24}{n^2}\frac{\kappa_{13}\kappa_{22}}{\kappa_{11}^3\kappa_{02}}+\frac{3}{n^2}\frac{\kappa_{22}^2}{\kappa_{11}^4}$$

$$+\frac{3}{n^2}\frac{\kappa_{20}^2\kappa_{02}^2}{\kappa_{11}^4}+\frac{6}{n^2}\frac{\kappa_{22}\kappa_{20}\kappa_{02}}{\kappa_{11}^4}\Big\}. \tag{6}$$

3. For the case when the parent population is the normal bivariate surface

$$\phi(xy)=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right],$$

$$\kappa_{10}=\kappa_{01}=0,\quad \kappa_{20}=\kappa_{02}=1,\quad \kappa_{11}=\rho \quad \text{and} \quad \kappa_{pq}=0 \quad \text{for} \quad p+q>2,$$

and the expressions obtained for the first four moments of r reduce to

$$\mu_1'(r)=\rho\left\{1-\frac{1}{2n}-\frac{3}{8n^2}+\rho^2\left(\frac{1}{2n}-\frac{3}{4n^2}\right)+\rho^4\frac{9}{8n^2}\right\}+\ldots, \tag{7}$$

$$\mu_2'(r)=\rho^2\left\{1-\frac{3}{n}+\frac{3}{n^2}+\rho^2\left(\frac{2}{n}-\frac{12}{n^2}\right)+\frac{1}{\rho^2}\left(\frac{1}{n}+\frac{1}{n^2}\right)+\rho^4\frac{8}{n^2}\right\}+\ldots, \tag{8}$$

$$\mu_3'(r)=\rho^3\left\{\frac{1}{\rho^2}\left(\frac{3}{n}-\frac{9}{2n^2}\right)+1-\frac{15}{2n}+\frac{261}{8n^2}+\rho^2\left(\frac{9}{2n}-\frac{225}{4n^2}\right)+\frac{225}{8n^2}\rho^4\right\}+\ldots, \tag{9}$$

$$\mu_4'(r)=\rho^4\left\{\frac{1}{\rho^4}\frac{3}{n^2}+\frac{1}{\rho^2}\left(\frac{6}{n}-\frac{36}{n^2}\right)+1-\frac{14}{n}+\frac{129}{n^2}+\rho^4\left(\frac{8}{n}-\frac{168}{n^2}\right)+\frac{72}{n^2}\rho^4\right\}+\ldots, \tag{10}$$

from which we deduce

$$\sigma^2(r)=\frac{1}{n}(1-\rho^2)^2\left(1+\frac{1}{n}+\frac{11\rho^2}{2n}\right), \tag{11}$$

$$\mu_3(r)=-\frac{6\rho}{n^2}(1-\rho^2)^3, \tag{12}$$

$$\mu_4(r)=\frac{3}{n^2}(1-\rho^2)^4. \tag{13}$$

The expressions for μ_2' and μ_4' correspond exactly (so far as order n^{-2}) to those given by Soper *et al.* (1917), which provides a check on the work.

4. To take formulae (3)–(6) any further (e.g. by using them in deriving a Gram-Charlier expansion for the frequency function of ρ) it would be necessary to make some simplifying assumptions concerning the cumulants. The difficulty is, that in doing so, one imposes such constraints on the distribution that it becomes doubtful what significance should be attributed to the coefficient r. In order to illustrate how the derived moments may be used it has been supposed that the parent population generating the sample is a Gram-Charlier surface of

order 3 only, with linear regression and equal marginal standard deviations. It is not claimed that this restricted case will be met with in practice. Writing r_N to represent the normal parent result

$$\mu_1'(r) = \mu_1'(r_N) + \rho\left\{(\kappa_{30}^2+\kappa_{03}^2)(1-\rho^2)\left(-\frac{3}{8n^2}+\frac{39}{16n^3}-\frac{45}{16n^3}\rho^2\right)\right.$$
$$\left.+\frac{\kappa_{30}\kappa_{03}}{\rho}(1-\rho^2)\left(-\frac{1}{4n^2}+\frac{9}{8n^3}+\frac{45}{8n^3}\rho^2\right)\right\}+\ldots, \quad (14)$$

$$\mu_2(r) = \mu_2(r_N) + \frac{1}{n^2}\rho^2(1-\rho^2)\left\{(\kappa_{30}^2+\kappa_{03}^2)\left(-\frac{5}{4}+\frac{59}{4n}\right)\right.$$
$$\left.+\frac{\kappa_{30}\kappa_{03}}{\rho}\left(\frac{5}{2}-\frac{57}{2n}+\frac{17\rho^2}{n}\right)\right\}+\frac{18\rho^6}{n^3}\left(\kappa_{30}^2+\kappa_{03}^2-2\frac{\kappa_{30}\kappa_{03}}{\rho}\right)+\ldots, \quad (15)$$

$$\mu_3(r) = \mu_3(r_N) + \frac{1}{n^2}\rho(1-\rho^2)\left\{(\kappa_{30}^2+\kappa_{03}^2)\left(\frac{9\rho^2}{4}-\frac{99\rho^2}{8n}-\frac{39\rho^4}{8n}\right)\right.$$
$$\left.+\frac{\kappa_{30}\kappa_{03}}{\rho}\left(-\frac{3\rho^2}{2}+\frac{3}{n}-\frac{75\rho^2}{4n}+\frac{309\rho^4}{4n}\right)\right\}+\frac{3\rho^5}{n^2}\left(1+\frac{9\rho^2}{n}\right)\left(\kappa_{30}^2+\kappa_{03}^2-2\frac{\kappa_{30}\kappa_{03}}{\rho}\right)+\ldots, \quad (16)$$

$$\mu_4(r) = \mu_4(r_N) + \frac{1}{n^3}\rho^2(1-\rho^2)\left\{(\kappa_{30}^2+\kappa_{03}^2)\left(-\frac{15}{2}-\frac{78}{2}\rho^2-\frac{51}{2}\rho^4\right)\right.$$
$$\left.+\frac{\kappa_{30}\kappa_{03}}{\rho}(15+6\rho^2+123\rho^4)\right\}+\ldots. \quad (17)$$

If the departure from normality is not too marked, then κ_{30} and κ_{03} will not be large and these expressions should give reasonable approximations to the true moments of r. The hypothesis which, in practice, it is usually desired to test is that $\rho = 0{\cdot}0$, that is, we are interested in finding out whether there is *some* association between our two variables. For this case the expressions for the moments simplify considerably and we have

$$\mu_1'(r) = \kappa_{30}\kappa_{03}\left(-\frac{1}{4n^2}+\frac{9}{8n^3}\right)+\ldots, \quad (18)$$

$$\mu_2(r) = \frac{1}{n-1}+\ldots, \quad (19)$$

$$\mu_3(r) = \frac{3\kappa_{30}\kappa_{03}}{n^3}+\ldots, \quad (20)$$

$$\mu_4(r) = \frac{3}{n^2-1}+\ldots, \quad (21)$$

giving

$$\beta_1(r) = \left(1-\frac{1}{n}\right)^3\frac{9\kappa_{30}^2\kappa_{03}^2}{n^3}+\ldots, \quad (22)$$

$$\beta_2(r) = \frac{3(n-1)}{(n+1)}+\ldots. \quad (23)$$

In this case we note that, to the degree of approximation considered, the amount by which the parent population deviates from normality does not affect the value of $\beta_2(r)$, and that, for reasonable-sized n, $\beta_1(r)$ does not differ appreciably from zero. Some values of $\beta_1(r)$ have been calculated and are given in Table 1. The cases are restricted to those for which $\kappa_{30}^2 = \kappa_{03}^2$.

It would thus appear that the same Pearson curve which represents the distribution of r in samples from a normal parent population will also approximate to the distribution of r when the parent population deviates slightly from the normal. Hence a small amount of skewness in the parent population will not affect the test of the hypothesis that $\rho = 0{\cdot}0$.

Table 1. *Values of $\beta_1(r)$ when parent population is not normal, $\rho = 0{\cdot}0$ and $\kappa_{30}^2 = \kappa_{03}^2$.*

n \ κ_{30}^2	0·0	0·2	0·4	0·6
20	0·0000	0·0000	0·0001	0·0003
25	0·0000	0·0000	0·0001	0·0002
50	0·0000	0·0000	0·0000	0·0000

5. As far as is known the distribution of the regression coefficients in samples from a non-normal population has not yet been investigated. If the regression of y on x is linear and, in the population, is of the form

$$y = \alpha + \beta x,$$

then we may write the regression line of y on x in the sample as

$$y = a + bx,$$

where the regression coefficient $\quad b = \dfrac{m_{11}}{m_{20}} = \dfrac{k_{11}}{k_{20}}.$

Writing b in the form of k-statistics, we have for the regression coefficient of y on x

$$b_{y|x} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{k_{11}}{k_{20}} = \frac{\kappa_{11}}{\kappa_{20}}\left(1+\frac{k_{11}-\kappa_{11}}{\kappa_{11}}\right)\left(1+\frac{k_{20}-\kappa_{20}}{\kappa_{20}}\right)^{-1}, \tag{24}$$

which can be expanded as a negative binomial for reasonably large n. Proceeding in a manner analogous to that used above to obtain the moments of the correlation coefficient, we have to order n^{-2}

$$\mu_1'(b_{y|x}) = \frac{\kappa_{11}}{\kappa_{20}}\left\{1+\left(\frac{1}{n}-\frac{2}{n^2}+\frac{3}{n^2}\frac{\kappa_{40}}{\kappa_{20}^2}\right)\left(\frac{\kappa_{40}}{\kappa_{20}^2}-\frac{\kappa_{31}}{\kappa_{20}\kappa_{11}}\right)-\frac{4}{n^2}\left(\frac{\kappa_{30}^2}{\kappa_{20}^3}-\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}\right)-\frac{1}{n^2}\left(\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}\right)\right\}, \tag{25}$$

$$\begin{aligned}\mu_2'(b_{y|x}) = \frac{\kappa_{11}^2}{\kappa_{20}^2}\Bigg\{&1-\frac{1}{n}-\frac{3}{n^2}+\frac{\kappa_{40}}{\kappa_{20}^2}\left(\frac{3}{n}-\frac{9}{n_2}\right)-\frac{4}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{16}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}+\frac{15}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}-\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}}\left(\frac{4}{n}-\frac{12}{n^2}\right)+\frac{6}{n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}\\&+\frac{24}{n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}-\frac{24}{n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}+\frac{\kappa_{22}}{\kappa_{11}^2}\left(\frac{1}{n}-\frac{4}{n^2}\right)+\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}\left(\frac{1}{n}+\frac{3}{n^2}\right)-\frac{2}{n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}+\frac{1}{n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}\\&-\frac{4}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}}-\frac{4}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}}+\frac{3}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}+\frac{6}{n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}^2}\Bigg\},\end{aligned} \tag{26}$$

$$\mu_3'(b_{y|x}) = \frac{\kappa_{11}^3}{\kappa_{20}^3}\left(1-\frac{3}{n}-\frac{9}{n^2}+\frac{\kappa_{40}}{\kappa_{20}^2}\left(\frac{6}{n}-\frac{30}{n^2}\right)+\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}}\left(-\frac{9}{n}+\frac{45}{n^2}\right)+\frac{\kappa_{22}}{\kappa_{11}^2}\left(\frac{3}{n}-\frac{21}{n^2}\right)+\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}\left(\frac{3}{n}+\frac{9}{n^2}\right)\right.$$
$$-\frac{10}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{40}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}+\frac{18}{n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}+\frac{72}{n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}-\frac{9}{n_2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}+\frac{9}{n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}-\frac{18}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}}$$
$$-\frac{18}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}}+\frac{1}{n^2}\frac{\kappa_{33}}{\kappa_{11}^3}+\frac{3}{n^2}\frac{\kappa_{20}\kappa_{13}}{\kappa_{11}^3}-\frac{6}{n^2}\frac{\kappa_{02}\kappa_{31}}{\kappa_{11}^3}+\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}^3}\frac{1}{n^2}+\frac{3}{n^2}\frac{\kappa_{12}\kappa_{21}}{\kappa_{11}^3}+\frac{45}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}$$
$$\left.-\frac{90}{n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}+\frac{36}{n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}^2}+\frac{18}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}-\frac{9}{n^2}\frac{\kappa_{31}\kappa_{22}}{\kappa_{11}^3\kappa_{20}}\right), \tag{27}$$

$$\mu_4'(b_{y|x}) = \frac{\kappa_{11}^4}{\kappa_{20}^4}\left(1-\frac{6}{n}-\frac{15}{n^2}+\frac{\kappa_{40}}{\kappa_{20}^2}\left(\frac{10}{n}-\frac{80}{n^2}\right)+\frac{\kappa_{31}}{\kappa_{11}\kappa_{20}}\left(-\frac{16}{n}+\frac{128}{n^2}\right)+\frac{\kappa_{22}}{\kappa_{11}^2}\left(\frac{6}{n}-\frac{66}{n^2}\right)\right.$$
$$+\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}\left(\frac{6}{n}+\frac{12}{n^2}\right)+\frac{105}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}-\frac{240}{n^2}\frac{\kappa_{40}\kappa_{31}}{\kappa_{11}\kappa_{20}^3}+\frac{60}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}+\frac{36}{n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}+\frac{120}{n^2}\frac{\kappa_{31}^2}{\kappa_{11}^2\kappa_{20}^2}-\frac{48}{n^2}\frac{\kappa_{31}\kappa_{22}}{\kappa_{11}^3\kappa_{20}}$$
$$-\frac{36}{n_2}\frac{\kappa_{31}\kappa_{02}}{\kappa_{11}^3}+\frac{3}{n^2}\frac{\kappa_{22}^2}{\kappa_{11}^4}+\frac{3}{n^2}\frac{\kappa_{20}^2\kappa_{02}^2}{\kappa_{11}^4}+\frac{6}{n^2}\frac{\kappa_{22}\kappa_{20}\kappa_{02}}{\kappa_{11}^4}-\frac{20}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3}-\frac{80}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}+\frac{40}{n^2}\frac{\kappa_{51}}{\kappa_{11}\kappa_{20}^2}+\frac{160}{n^2}\frac{\kappa_{30}\kappa_{21}}{\kappa_{11}\kappa_{20}^2}$$
$$\left.-\frac{24}{n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}-\frac{48}{n^2}\frac{\kappa_{30}\kappa_{12}}{\kappa_{11}^2\kappa_{20}}-\frac{48}{n^2}\frac{\kappa_{21}^2}{\kappa_{11}^2\kappa_{20}}+\frac{4}{n^2}\frac{\kappa_{33}}{\kappa_{11}^3}+\frac{12}{n^2}\frac{\kappa_{20}\kappa_{13}}{\kappa_{11}^3}+\frac{4}{n^2}\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}^3}+\frac{12}{n^2}\frac{\kappa_{21}\kappa_{12}}{\kappa_{11}^3}\right). \tag{28}$$

6. If we assume that the regressions of x on y and y on x are linear, then

$$\kappa_{20}\kappa_{p1} = \kappa_{11}\kappa_{p+1,0}, \quad \kappa_{02}\kappa_{1p} = \kappa_{11}\kappa_{0,p+1}, \quad \text{where} \quad p \geqslant 1, \tag{29}$$

and the expressions for the first four moments of $b_{y|x}$ reduce to

$$\mu_1'(b_{y|x}) = \frac{\kappa_{11}}{\kappa_{20}}, \tag{30}$$

$$\mu_2(b_{y|x}) = \frac{\kappa_{11}^2}{\kappa_{20}^2}\left(-\frac{1}{n}-\frac{3}{n^2}-\frac{\kappa_{40}}{\kappa_{20}^2}\left(\frac{1}{n}-\frac{3}{n^2}\right)+\frac{2}{n^2}\frac{\kappa_{60}}{\kappa_{20}^3}+\frac{4}{n^2}\frac{\kappa_{30}^2}{\kappa_{20}^3}-\frac{3}{n^2}\frac{\kappa_{40}^2}{\kappa_{20}^4}+\frac{\kappa_{22}}{\kappa_{11}^2}\left(\frac{1}{n}-\frac{4}{n^2}\right)\right.$$
$$\left.+\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}\left(\frac{1}{n}+\frac{3}{n^2}\right)-\frac{2}{n^2}\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}+\frac{1}{n^2}\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}-\frac{4}{n^2}\frac{\kappa_{30}\kappa_{03}}{\kappa_{20}\kappa_{11}\kappa_{02}}+\frac{3}{n^2}\frac{\kappa_{40}\kappa_{22}}{\kappa_{20}^2\kappa_{11}^2}\right), \tag{31}$$

$$\mu_3(b_{y|x}) = \frac{\kappa_{11}^3}{\kappa_{20}^3}\frac{1}{n^2}\left(6\frac{\kappa_{40}}{\kappa_{20}^2}-9\frac{\kappa_{22}}{\kappa_{11}^2}+2\frac{\kappa_{60}}{\kappa_{20}^3}+2\frac{\kappa_{30}^2}{\kappa_{20}^3}-3\frac{\kappa_{42}}{\kappa_{11}^2\kappa_{20}}+\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}^3}+\frac{\kappa_{33}}{\kappa_{11}^3}+3\frac{\kappa_{20}\kappa_{04}}{\kappa_{11}^2\kappa_{02}}-3\frac{\kappa_{30}\kappa_{03}}{\kappa_{11}\kappa_{20}\kappa_{02}}\right), \tag{32}$$

$$\mu_4(b_{y|x}) = \frac{\kappa_{11}^4}{\kappa_{20}^4}\frac{1}{n^2}\left(3+6\frac{\kappa_{40}}{\kappa_{20}^2}-6\frac{\kappa_{22}}{\kappa_{11}^2}-6\frac{\kappa_{20}\kappa_{02}}{\kappa_{11}^2}+3\frac{\kappa_{40}^2}{\kappa_{20}^4}-6\frac{\kappa_{40}\kappa_{22}}{\kappa_{11}^2\kappa_{20}^2}\right.$$
$$\left.+3\frac{\kappa_{22}^2}{\kappa_{11}^4}-6\frac{\kappa_{40}\kappa_{02}}{\kappa_{11}^2\kappa_{20}}+3\frac{\kappa_{20}^2\kappa_{02}^2}{\kappa_{11}^4}+6\frac{\kappa_{22}\kappa_{20}\kappa_{02}}{\kappa_{11}^4}\right), \tag{33}$$

to order n^{-2}.

When the parent population is a bivariate normal distribution, we have further

$$\mu_1'(b_{y|x}) = \frac{\kappa_{11}}{\kappa_{20}} = \rho\frac{\sigma_y}{\sigma_x}, \tag{34}$$

$$\mu_2(b_{y|x}) = \left(\frac{1}{n}+\frac{3}{n^2}\right)\left(\frac{\kappa_{02}}{\kappa_{20}}-\frac{\kappa_{11}^2}{\kappa_{20}^2}\right) = \left(\frac{1}{n}+\frac{3}{n^2}\right)(1-\rho^2)\frac{\sigma_y^2}{\sigma_x^2}, \tag{35}$$

$$\mu_3(b_{y|x}) = 0, \tag{36}$$

$$\mu_4(b_{y|x}) = \frac{3}{n^2}\left(\frac{\kappa_{02}}{\kappa_{20}}-\frac{\kappa_{11}^2}{\kappa_{20}^2}\right)^2 = \frac{3}{n^2}(1-\rho^2)^2\frac{\sigma_y^4}{\sigma_x^4}, \tag{37}$$

which agree to our order of approximation with the known moments of the distribution of $b_{y|x}$ for a normal distribution.

Table 2. *Values of* $\beta_1(b)$, $\beta_2(b)$ *for* $\rho = 0{\cdot}0$, $\kappa_{30}^2 = \kappa_{03}^2$

n	20		40		60	
κ_{30}^2	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$
0·0	0·0000	3·4000	0·0000	3·1714	0·0000	3·1091
0·2	0·0002	3·3827	0·0000	3·1279	0·0000	3·1071
0·4	0·0010	3·3656	0·0001	3·0852	0·0000	3·1052

Table 3. *Values of* $\beta_1(b)$, $\beta_2(b)$ *for* $\rho \neq 0{\cdot}0$, $\kappa_{30}^2 = \kappa_{03}^2$

(*a*) $\rho = 0{\cdot}2$, $\kappa_{30} = \kappa_{03}$

n	20		40		60	
κ_{30}^2	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$
0·0	0·0000	3·4000	0·0000	3·1714	0·0000	3·1091
0·2	0·0000	3·3999	0·0000	3·1746	0·0000	3·1120
0·4	0·0001	3·3996	0·0001	3·1777	0·0000	3·1148
0·6	0·0019	3·3990	0·0003	3·1808	0·0001	3·1176

(*b*) $\rho = 0{\cdot}2$, $\kappa_{30} = -\kappa_{03}$

n	20		40		60	
κ_{30}^2	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$
0·0	0·0000	3·4000	0·0000	3·1714	0·0000	3·1091
0·2	0·0000	3·3573	0·0000	3·1559	0·0000	3·1001
0·4	0·0004	3·3141	0·0001	3·1401	0·0000	3·0911
0·6	0·0010	3·2706	0·0002	3·1242	0·0000	3·0817

(*c*) $\rho = 0{\cdot}4$, $\kappa_{30} = \kappa_{03}$

n	20		40		60	
κ_{30}^2	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$	$\beta_1(b)$	$\beta_2(b)$
0·0	0·0000	3·4000	0·0000	3·1714	0·0000	3·1091
0·2	0·0000	3·4062	0·0000	3·1765	0·0000	3·1130
0·4	0·0001	3·4118	0·0000	3·1809	0·0000	3·1167
0·6	0·0003	3·4168	0·0000	3·1861	0·0000	3·1204

7. For the same double Gram-Charlier surface as before (linear regressions and equal marginal standard deviations), we have for the moments

$$\mu_1'(b_{y|x}) = \rho, \tag{38}$$

$$\mu_2(b_{y|x}) = \frac{1-\rho^2}{n-3} + \frac{2\rho\kappa_{30}}{n^2}(\kappa_{03}-\rho\kappa_{30})\left(2-\frac{15}{n}\right) + \frac{6}{n^3}\kappa_{30}(\kappa_{30}-\rho\kappa_{03}), \tag{39}$$

$$\mu_3(b_{y|x}) = \frac{9\rho}{n^3}\kappa_{30}(\kappa_{30}-\rho\kappa_{03}) + \frac{9}{n^3}\kappa_{30}(1-2\rho^2)(\kappa_{03}-\rho\kappa_{30}), \tag{40}$$

$$\mu_4(b_{y|x}) = \frac{3(1-\rho^2)^2}{(n-3)(n-5)} + \frac{12\rho^2}{n^3}(\kappa_{03}-\rho\kappa_{30})^2 + \frac{36\rho}{n^3}(\kappa_{30}-\rho\kappa_{03})(\kappa_{03}-\rho\kappa_{30}). \tag{41}$$

The β_1 and β_2 of the distribution of $b_{y|x}$ have been calculated from the above expressions for various values of ρ and κ_{03}, κ_{30}, and it appears that a small amount of skewness will not seriously affect the distribution of b. Tables 2 and 3 show some values of β_1 and β_2 for certain values of ρ, κ_{30} and κ_{03}.

8. It is clear from a study of these tables that we might expect the effect of skewness in the parent population on the sampling distribution of $b_{y|x}$ to be slight. However, we may really only judge this effect by approximating to the distribution of $b_{y|x}$ in some way (for example by a Pearson curve with the correct first four moments) and finding out whether the nominal significance level is affected to any appreciable extent. This procedure was carried out for each of the above cases considered, and the results show definitely that nowhere were the 5 and 1 % significance levels altered enough to invalidate any test of significance carried out on the assumption of parent normality. We may therefore reach the conclusion that distortion in the way described above of the normal parent population is unlikely to lead to error in tests of significance.

I should like to thank Dr F. N. David and Prof. M. G. Kendall for helpful suggestions in preparing this paper for press.

REFERENCES

Cook, M. B. (1951). *Biometrika*, **38**, 179.

Soper, H. E., Young, A. W., Cave, B. M., Lee, A. & Pearson, K. (1917). *Biometrika*, **11**, 328.

INTERRELATIONS BETWEEN CERTAIN LINEAR SYSTEMATIC STATISTICS OF SAMPLES FROM ANY CONTINUOUS POPULATION

By G. P. SILLITTO

Research Department, Imperial Chemical Industries Ltd., Nobel Division

1. Statistics which consist of linear combinations of the observations in a sample have been called by Mosteller (1946) 'systematic statistics'. In one class of these the law of combination of the observations is always the same for samples of a given size. It is evident, therefore, that linear systematic statistics of this class must be linear combinations of each other. Their expectations will be linearly related, and the variance of one will be expressible in terms of variances and co-variances of others. These relations will hold for samples from all continuous populations. In this paper, expressions are obtained for the expectations of some linear systematic statistics of this class in terms of expectations of others.

It is found that certain linear systematic statistics which measure dispersion, such as Gini's coefficient of mean difference, have expectations which are independent of the size of the sample from which they are calculated, and linear systematic statistics having this property are derived which may be used as indices of skewness and kurtosis for samples from any continuous population. Statistics with this property have advantages, for instance when estimates from samples of different sizes are to be combined.

2. The expectation of the range, $\overline{w}_n$, in a sample of n from any continuous population of variate-values z, which has distribution function F, is (cf. Kendall, 1943, p. 223)

$$\overline{w}_n = \int \{1-(1-F)^n - F^n\}\, dz. \tag{1}$$

Similarly,
$$\overline{w}_{n-1} = \int \{1-(1-F)^{n-1} - F^{n-1}\}\, dz,$$

the integrals being taken over the whole range of z. Therefore

$$\begin{aligned}\overline{w}_n &= \int \{(1-F)^{n-1} F + F^{n-1}(1-F)\}\, dz + \overline{w}_{n-1} \\ &= (\chi_{n,1} + \chi_{n,n-1})/n + \overline{w}_{n-1}, \end{aligned} \tag{2}$$

where $\chi_{n,p}$ is the expectation of the difference between the $(p+1)$th and the pth value of z when the sample members are arranged in ascending order of magnitude. It is known (K. Pearson, 1902) that

$$\chi_{n,p} = \binom{n}{p} \int F^{n-p}(1-F)^p\, dz. \tag{3}$$

Obviously $\overline{w}_{n-1}$ in (2) can be expressed in terms of $\chi_{n-1,1}, \chi_{n-1,n-2}$ and $\overline{w}_{n-2}$, and so on.

3. From (3)

$$\begin{aligned}\frac{n}{p}\chi_{n-1,p-1} - \frac{n-p+1}{p}\chi_{n,p-1} &= \int \left\{ \binom{n}{p} F^{n-p}(1-F)^{p-1} - \binom{n}{p} F^{n-p+1}(1-F)^{p-1} \right\} dz \\ &= \binom{n}{p} \int F^{n-p}(1-F)^p\, dz = \chi_{n,p}, \end{aligned} \tag{4}$$

and by using this relation repeatedly on the terms of its own left-hand side it is found that

$$\chi_{n,p} = \frac{n!\,(p-\nu)!}{(n-\nu)!\,p!}\left\{\chi_{n-\nu,p-\nu} - \nu\frac{n-p+1}{n-\nu+1}\chi_{n-\nu+1,p-\nu} + \binom{\nu}{2}\frac{(n-p+1)(n-p+2)}{(n-\nu+1)(n-\nu+2)}\chi_{n-\nu+2,p-\nu} - \dots\right.$$
$$\left. + (-1)^{\nu}\frac{(n-p+1)(n-p+2)\dots(n-p+\nu)}{(n-\nu+1)(n-\nu+2)\dots(n-1)\,n}\chi_{n,p-\nu}\right\}. \tag{5}$$

In the special case of $\nu = p-1$, (5) becomes

$$\chi_{n,p} = \binom{n}{p}\left\{\frac{\chi_{n-p+1,1}}{n-p+1} - \binom{p-1}{1}\frac{\chi_{n-p+2,1}}{n-p+2} + \binom{p-1}{2}\frac{\chi_{n-p+3,1}}{n-p+3} - \dots + (-1)^{p-1}\frac{\chi_{n,1}}{n}\right\}. \tag{6}$$

4. By using (4) in the right-hand side of relations like (2), other expressions for $\overline{w}_n$ in terms of mean differences in samples of size n and smaller can be obtained; and from such expressions one can also express $\chi_{n,p} + \chi_{n,n-p}$ in terms of mean ranges in samples of size n and less. The relation obtained is proved quite briefly in the following way:

$$-\sum_{r=0}^{p}(-1)^r\binom{p}{r}\overline{w}_{n-p+r} = \int -\left\{\Sigma(-1)^r\binom{p}{r} - \Sigma(-1)^r\binom{p}{r}(1-F)^{n-p+r} - \Sigma(-1)^r\binom{p}{r}F^{n-p+r}\right\}dz$$
$$= \int -\{(1-1)^p - (1-F)^{n-p}F^p - F^{n-p}(1-F)^p\}\,dz$$
$$= \frac{p!\,(n-p)!}{n!}(\chi_{n,p} + \chi_{n,n-p}). \tag{7}$$

The form of this expression which arises in a symmetrical population, in which $\chi_{n,p} = \chi_{n,n-p}$, was proved by E. S. Pearson (1926).

For n odd, we have from (7), taking p first equal to 1 and secondly to $n-1$,

$$\frac{1!\,(n-1)!}{n!}(\chi_{n,1} + \chi_{n,n-1}) = \overline{w}_n - \overline{w}_{n-1},$$

$$\frac{(n-1)!\,1!}{n!}(\chi_{n,n-1} + \chi_{n,1}) = -\overline{w}_n + \binom{n-1}{1}\overline{w}_{n-1} - \binom{n-1}{2}\overline{w}_{n-2} + \dots + \binom{n-1}{1}\overline{w}_2.$$

Therefore
$$\overline{w}_n = \frac{1}{2}\left\{\binom{n-1}{1}\overline{w}_{n-1} - \binom{n-1}{2}\overline{w}_{n-2} + \dots + \binom{n-1}{1}\overline{w}_2\right\} + \tfrac{1}{2}\overline{w}_{n-1}, \tag{8}$$

and similarly a number of relations between mean ranges can be obtained by giving other values to p. By means of (7) the relation between mean ranges derived by Romanowsky (1933) for a normal population can be proved for any continuous population.

A special case of (8) pointed out by Robbins (1944) is

$$\overline{w}_3 = \tfrac{3}{2}\overline{w}_2. \tag{9}$$

5. K. R. Nair (1950) has proposed the symbol $j_{(r)}$ for the difference between the sum of the largest r and the smallest r members of a sample of n ($1 \leqslant r \leqslant \frac{1}{2}n$ or $\frac{1}{2}(n-1)$ according as n is even or odd) which was suggested by Jones (1946) as a measure of dispersion in large samples. We shall use ${}_nj_{(r)}$ to denote this statistic in a sample of n, so that

$${}_nj_{(r)} = (z_n + z_{n-1} + \dots + z_{n-r+1}) - (z_r + z_{r-1} + \dots + z_1)$$
$$= (z_n - z_1) + (z_{n-1} - z_2) + \dots + (z_{n-r+1} - z_r).$$

Therefore

$$E({}_nj_{(r)}) = \overline{w}_n + (\overline{w}_n - \chi_{n,n-1} - \chi_{n,1}) + \ldots + \{\overline{w}_n - (\chi_{n,n-1} + \chi_{n,1}) - (\chi_{n,n-2} + \chi_{n,2}) - \ldots - (\chi_{n,n-r+1} + \chi_{n,r-1})\}. \quad (10)$$

By means of (7) the right-hand side can be expressed as a series of mean ranges. The coefficients of each of the $\overline{w}_{n-s}$ $(0 \leqslant s \leqslant r-1)$ can then be summed, giving

$$E({}_nj_{(r)}) = \sum_{s=0}^{r-1} (-1)^{r-s-1} \binom{n-s-2}{r-s-1} \binom{n}{s} \overline{w}_{n-s}. \quad (11)$$

6. If $x_{n,p}$ is the pth difference between the ranked individuals in a sample of n, the mean deviation from the median in the sample is, if n is odd,

$$m'_n = \frac{1}{n}\{x_{n,1} + 2x_{n,2} + \ldots + \tfrac{1}{2}(n-1)\, x_{n,\frac{1}{2}(n-1)} + \tfrac{1}{2}(n-1)\, x_{n,\frac{1}{2}(n+1)} + \ldots + 2x_{n,n-2} + x_{n,n-1}\},$$

if n is even

$$m'_n = \frac{1}{n}\{x_{n,1} + 2x_{n,2} + \ldots + (\tfrac{1}{2}n-1)\, x_{n,\frac{1}{2}n-1} + \tfrac{1}{2}n x_{n,\frac{1}{2}n} + (\tfrac{1}{2}n-1)\, x_{n,\frac{1}{2}n+1} + \ldots + 2x_{n,n-2} + x_{n,n-1}\}.$$

By taking expectations and using (4), it is easily found that

$$E(m'_{2k-1}) = E(m'_{2k}) - \frac{1}{2(2k-1)} \chi_{2k,k}, \quad (12)$$

and that

$$E(m'_{2k+1}) = E(m'_{2k}). \quad (13)$$

The latter relation has also been proved by Godwin (1949). Thus we obtain

$$E(m'_{2k+1}) = E(m'_{2k}) = \frac{1}{2} \sum_{r=1}^{k} \frac{1}{2r-1} \chi_{2r,r}, \quad (14)$$

which can also be expressed in terms of mean ranges by use of (7).

7. From the $n-1$ expressions obtainable from (7) by giving p the values $1, 2, \ldots, n-1$ there arise $n-1$ independent equations relating $\overline{w}_2, \overline{w}_3, \ldots, \overline{w}_n$ and $\chi_{n,1}, \chi_{n,2}, \ldots, \chi_{n,n-1}$. These equations can then be solved to derive explicit expressions for $\overline{w}_m$ $(2 \leqslant m \leqslant n)$ in terms of mean differences in a sample of size n. The general expression obtained is

$$\overline{w}_m = \sum_{p=1}^{n-1} \chi_{n,p} - \frac{(n-m)!}{n!} \sum_{p=1}^{n-m} \frac{(n-p)!}{(n-m-p)!} (\chi_{n,p} + \chi_{n,n-p}), \quad (15)$$

which can be verified directly from the integral expressions for $\overline{w}_m$ and $\chi_{n,p}$. Relation (15) was obtained in another way by E. S. Pearson (1926).

If in (15) we put $m = 2$, we obtain after some reduction

$$\overline{w}_2 = \frac{2}{n(n-1)}\{(n-1)\chi_{n,1} + 2(n-2)\chi_{n,2} + \ldots + p(n-p)\chi_{n,p} + \ldots + (n-1)\chi_{n,n-1}\}. \quad (16)$$

The right-hand side can be recognized as the form taken by the expectation of g, Gini's coefficient of mean difference (Gini, 1912), when this is expressed in terms of successive differences of the ranked variate-values.

When the values $x_{n,p}$ of the pth successive difference ($p = 1, 2, \ldots, n-1$) between the ranked members of a sample of n observations are inserted on the right-hand side of (15) instead of their expectations $\chi_{n,p}$, the quantities so obtained may be regarded as statistics which are generalizations of Gini's coefficient of mean difference. We may refer to them as the g_m statistics, such that $E(g_m) = \overline{w}_m$ for samples of any size $n \geqslant m$, from any continuous population; thus

$$g_m = \sum_{p=1}^{n-1} x_{n,p} - \frac{(n-m)!}{n!} \sum_{p=1}^{n-m} \frac{(n-p)!}{(n-m-p)!} (x_{n,p} + x_{n,n-p}). \tag{17}$$

The expectations of the g_m statistics are independent of the number of observations in the sample from which they are calculated, unlike Jones's $j_{(r)}$ statistics, or the mean deviation from the median.

Table 1. *Variance of estimates of $\overline{w}_m$ from a sample of size n by methods (i), (ii), (iii) and (iv)*

m	Method	Sample size n				
		4	6	8	9	10
2	(i)	0·230	0·135	0·096	0·083	0·074
	(ii)	0·363	0·242	0·182	—	0·145
	(iii)	0·233	0·143	0·106	0·094	0·085
	(iv)	0·227	0·133	0·094	0·082	0·073
3	(i)	0·517	0·304	0·215	0·188	0·167
	(ii)	—	0·395	—	0·263	—
	(iii)	0·523	0·321	0·238	0·212	0·192
	(iv)	0·510	0·299	0·211	0·184	0·163
4	(i)	—	0·448	0·316	0·276	0·244
	(ii)	—	—	0·387	—	—
	(iii)	—	0·475	0·351	0·314	0·284
	(iv)	—	0·443	0·313	0·273	0·242
5	(i)	—	0·582	0·406	0·353	0·313
	(ii)	—	—	—	—	0·373
	(iii)	—	0·606	0·449	0·400	0·363
	(iv)	—	0·565	0·399	0·348	0·308
6	(i)	—	—	0·493	0·427	0·377
	(ii)	—	—	—	—	—
	(iii)	—	—	0·533	0·475	0·431
	(iv)	—	—	0·474	0·413	0·366

It is sometimes necessary to estimate the mean range in a sample of m from a sample of n ($> m$) observations; for instance, in quality control work or when it is proposed to use $\overline{w}_m$ as an index of dispersion. The g_m statistics could be used for this purpose, and it appeared of some interest to compare their efficiency as estimators of $\overline{w}_m$ in the case of a normal population, with some other methods. Table 1 shows the variances of estimates of $\overline{w}_m$, as obtained from a sample of n ($> m$) drawn from a normal population of unit standard deviation, by

(i) Using g_m as estimator. The variance of the estimate is obtainable from the data in Table 2 of Godwin's (1949) paper, since the estimator is a linear combination of the differences between ranked members of the sample.

(ii) Dividing the n observations into k subsamples of m and estimating $\overline{w}_m$ as the mean of the ranges in these subsamples. The variance of the estimate is $1/k$ times the variance of that from a single sample of m and can be calculated from Godwin's data.

(iii) Taking $\overline{w}_m$ as the observed range in the sample of $n \times d_m/d_n$. The variance of the estimate is $d_m^2 \operatorname{var} w_n/d_n^2$. ($d_n$ is the expectation of w/σ in a normal sample of n.)

(iv) Calculating the standard deviation s and estimating $\overline{w}_m$ as $s d_m$. Using the maximum-likelihood estimator of the standard deviation, the variance of the estimate of $\overline{w}_m$ is $d_m^2\left\{\frac{1}{2}(n-1)\left(\frac{\Gamma(\frac{1}{2}(n-1))}{\Gamma(\frac{1}{2}n)}\right)^2 - 1\right\}$.

K. R. Nair (1949) has investigated the efficiency of g (g_2 in this notation) as an index of dispersion in normal samples, in comparison with the linear systematic statistic which is most efficient as an index of dispersion (Godwin, 1949), and found it to be high.

8. As a linear measure of skewness it appears natural to consider the difference between the pth and the $(n-p)$th difference in a sample of n, or more generally a linear combination of such differences. It proves possible to choose a linear combination which, like the g_m statistics, has an expectation value independent of the size of the sample, and thus is specially suitable as a measure.

We seek an expression

$$ {}_n a_1(\chi_{n,1} - \chi_{n,n-1}) + {}_n a_2(\chi_{n,2} - \chi_{n,n-2}) + \ldots + {}_n a_p(\chi_{n,p} - \chi_{n,n-p}) + \ldots $$

(with $p = 1, 2, \ldots, \frac{1}{2}(n-1)$ if n is odd, and $p = 1, 2, \ldots, (\frac{1}{2}n - 1)$ if n is even) whose value is independent of the sample size. If it exists it must clearly equal ${}_3a_1(\chi_{3,1} - \chi_{3,2})$. If for all the permissible values of p we express $\chi_{n,p} - \chi_{n,n-p}$ in terms of first differences by using (6) and remember that

$$ \chi_{2,1} = \overline{w}_2 = \tfrac{2}{3}\overline{w}_3 = \tfrac{2}{3}(\chi_{3,1} + \chi_{3,2}), $$

the resulting equations can be solved for $\chi_{3,1} - \chi_{3,2}$. In fact we obtain

$$ \tfrac{1}{3}(\chi_{3,1} - \chi_{3,2}) = \sum_p \left[(\chi_{n,p} - \chi_{n,n-p}) \left\{ \binom{n-2}{p-1} \frac{n-2p}{n-2} \right\Big/ \binom{n}{p} \right], \tag{18} $$

the sum being taken over the permissible values of p. This may be verified from the integral expression for $\chi_{n,p}$.

Thus the statistic

$$ S_L = \sum_p \left[(x_{n,p} - x_{n,n-p}) \left\{ \binom{n-2}{p-1} \frac{n-2p}{n-2} \right\Big/ \binom{n}{p} \right] \tag{19} $$

is a linear systematic statistic which measures skewness in samples of n from any continuous population. Its expectation $\frac{1}{3}(\chi_{3,1} - \chi_{3,2})$ is the expectation of the difference (median-mean) in samples of three from the population. The more familiar measure due to K. Pearson, the ratio of (mean-mode) to standard deviation, is not independent of the population, which has to be known to determine the mode.

9. It does not appear possible to define kurtosis except by reference to some standard population. The virtue of the normal population for this purpose is not marked except when one uses measures derived from moments, and in considering continuous populations in general, the most natural one to take as a standard appears to be the rectangular, which is flat throughout its range. The smallest sample which could give an indication of the kurtosis is that of four values. In such a sample, an index of kurtosis relative to the rectangular

population would be $\frac{1}{2}(x_{4,1}-2x_{4,2}+x_{4,3})$, which has expected value zero in the rectangular population. One could therefore take as a measure of kurtosis a linear combination of the differences between the values in the sample, which has the same expectation as this expression, independently of the size of the sample.

It is easily seen that $\frac{1}{2}(\chi_{4,1}-2\chi_{4,2}+\chi_{4,3}) = 5\overline{w}_4-9\overline{w}_2$ and the mean differences in any sample of more than four can be expressed in terms of $\overline{w}_4$ and $\overline{w}_2$ by means of (15). Thus a linear systematic statistic for measuring kurtosis would be

$$K_L = 5g_4-9g_2. \tag{20}$$

In comparing values of this statistic (or of S_L) one would of course bring them to the same scale by dividing them by a measure of the dispersion such as the standard deviation, or $\overline{w}_2$ if one wishes to restrict oneself to linear statistics. The expectation of K_L in the case of the normal distribution, with unit standard deviation, is $+0{\cdot}1383$, or if the distribution is rescaled by dividing by $\overline{w}_2$ it is $+0{\cdot}1226$. As another example, the symmetrical U-shaped Type I distribution with frequency function

$$f(z) = \frac{1}{\pi\sqrt{[z(1-z)]}} \quad (0<z<1)$$

is easily found to have

$$\overline{w}_2 = \frac{4}{\pi^2}, \quad \overline{w}_4 = \frac{12}{\pi^2}\left(1-\frac{4}{\pi^2}\right), \quad \sigma = \frac{1}{2\sqrt{2}},$$

so that K_L has the expectation

$$6\left(1-\frac{10}{\pi^2}\right) = -0{\cdot}07927 \quad \text{or} \quad \frac{48\sqrt{2}}{\pi^2}\left(1-\frac{10}{\pi^2}\right) = -0{\cdot}09087$$

when scaled in terms of $\overline{w}_2$ or σ respectively.

The author is indebted to Prof. E. S. Pearson for helpful suggestions and advice in presenting this paper.

REFERENCES

Gini, C. (1912). *Variabilità e Mutabilità*. Studi Economico-Giuridico della Reale Università di Cagliari: Anno 3, part 2, p. 80.
Godwin, H. J. (1949). *Biometrika*, **36**, 92.
Jones, A. E. (1946). *Biometrika*, **33**, 274.
Kendall, M. G. (1943). *The Advanced Theory of Statistics*, **1**. London: Chas. Griffin and Co.
Mosteller, F. (1946). *Ann. Math. Statist.* **17**, 377.
Nair, K. R. (1949). *Bull. Calcutta Statist. Ass.* **2**, 129.
Nair, K. R. (1950). *Biometrika*, **37**, 182.
Pearson, E. S. (1926). *Biometrika*, **18**, 173.
Pearson, K. (1902). *Biometrika*, **1**, 385.
Robbins, H. E. (1944). *Ann. Math. Statist.* **15**, 323.
Romanowsky, V. (1933). *Biometrika*, **25**, 195.

THE DELAY TO PEDESTRIANS CROSSING A ROAD

By J. C. TANNER

Road Research Laboratory, Department of Scientific and Industrial Research, Harmondsworth

1. Introduction

Suppose a pedestrian is waiting to cross a road. How long may he expect to have to wait before there is a sufficient gap in the traffic? The answer will depend on a number of factors, but an answer can be given if we make certain reasonable assumptions about his behaviour and that of the traffic. The mathematical theory which arises could be stated in more general terms, but for simplicity the application to the pedestrian problem is considered throughout. It is thought, however, that useful applications for it may exist in other fields.

The simplest form of the process may be stated as follows: We have a road in which traffic is flowing past a point X, in one direction only, so that

(i) the number of arrivals in a time t follows a Poisson distribution with mean Nt, so that the probability that r vehicles pass in this time is

$$e^{-Nt}(Nt)^r/r! \quad (r = 0, 1, 2, \ldots);$$

(ii) in non-overlapping intervals of lengths t and t', the numbers of arrivals are statistically independent.

We assume, in other words, that vehicles arrive 'at random'. The assumption of randomness is not usually true in practice, but Adams (1936) has shown that the approximation is quite good under a fairly wide range of conditions. Pedestrians arrive at X at a rate n per unit time, also at random, and independently of the passage of vehicles. Each pedestrian starts to cross the road as soon as it can be seen that no vehicle will pass X during the next interval I, a time assumed for the moment not to vary from one pedestrian to another. In practice, pedestrians often arrive in twos and threes, these 'arrival groups' arriving sufficiently at random, and the theory will apply in this case to groups of pedestrians. However, some indication of how the theory can be applied to individuals rather than arrival groups is given.

Each pedestrian will be delayed for a time T, varying from one pedestrian to another. Let $P(T)$ be the probability of a delay greater than T, and let $f(T)$, when it exists, be the frequency function of T, so

$$f(T) = -\frac{dP(T)}{dT}.$$

We shall derive for the above hypothetical situation the function $P(T)$ and the distributions of the sizes of groups of pedestrians crossing or waiting at certain times.

Some of the results are extended to the cases when there are two streams of traffic, when the critical gap I is not constant, and when we are given the proportion of arrival-groups of different sizes. The theory has been applied to estimate the delays which would occur to pedestrians and vehicles if 'advance warning signs' were erected at pedestrian crossings. These are intended to give more guidance to road users as to who has the right of way.

The results are not all new; the simplest formula for the mean delay to pedestrians was quoted by Adams (1936). Garwood (1940) considered a problem connected with traffic lights which was mathematically almost the same as this, and later applied the results to pedestrian delays for one and two traffic streams. Since this paper was submitted for publication, Raff (1951) has derived the mean delay by a consideration of 'blocks' of vehicles, as in §2·3 of the present paper. He also notes equation (7) of §2·2·1.

2. The distribution of pedestrian delay (simplest case)

Three derivations will now be given of the function $P(T)$, already defined.

2·1. Garwood (1940) considered a very similar problem in connexion with the operation of vehicular-controlled traffic lights. In this, vehicles on the minor road arrive at random at a junction with a major road, and if the lights are against them, the lights change as soon as there has been a traffic-free gap I at the detector pads on the major road, or after a time M if this is sooner. Garwood gave the formula for the probability of a vehicle on the minor road being delayed for more than a time T. His derivation proceeded from an expression for the probability distribution of the longest interval in a line of unit length divided by n random points. It is easy to see that the probability $P(T)$ that a pedestrian is delayed for longer than T is the same as the probability that a vehicle on the minor road is delayed for more than $T+I$, for an infinite value of M, and substitution in his formula gives

$$P(T) = \sum_{s=0}^{r+1} \frac{(-)^s e^{-sNI} N^s (T-sI+I)^s}{s!} + \sum_{s=1}^{r+1} \frac{(-)^s e^{-sNI} N^{s-1} (T-sI+I)^{s-1}}{s-1!}, \tag{1}$$

where r is the integral part of T/I.

2·2. The next derivation of $P(T)$ is perhaps the most elegant, though it is possibly less useful than the third one. It is convenient to change the time-scale so that the unit of time is equal to the critical vehicle interval I. Let w and q refer to waiting time and flow on this scale, so that

$$w = T/I, \tag{2}$$

and

$$q = NI. \tag{3}$$

It is also convenient to define

$$a = qe^{-q}. \tag{4}$$

The symbols P and f will be retained, since this will not lead to any ambiguity.

2·2·1. A pedestrian crosses immediately if no vehicle is due to arrive in the next unit time, and the probability of this is e^{-q}. Hence

$$P(0) = 1 - e^{-q}. \tag{5}$$

If $0 < w < 1$, a pedestrian will start to cross in $(w, w+dw)$ if and only if (i) a vehicle has just passed, in $(w, w+dw)$; and (ii) no other vehicle will pass in $(w, w+1)$. One expression for the probability of this event is $-P'(w)\,dw$. (It is easy to see that $P(w)$ is differentiable for $0 < w < 1$.) It is also the product of the probabilities of (i) and (ii), which leads to

$$P'(w) = -a. \tag{6}$$

If $w > 1$, one other condition must be added to the above, namely, (iii) the pedestrian was waiting at time $w-1$. This is independent of the other two events, and has probability $P(w-1)$. Thus we have

$$P'(w) = -aP(w-1) \quad (w > 1). \tag{7}$$

At $w = 1$, $P'(w)$ is evidently not continuous, though $P(w)$ will be so. It is clear that in the ranges $(0, 1)$, $(1, 2)$, $(2, 3)$, ..., $P(w)$ is successively a linear, quadratic, ... function of w. By successive integrations the function $P(w)$ can be obtained as

$$P(w) = \theta(w) - e^{-q}\theta(w-1), \tag{8}$$

where

$$\theta(w) = 1 - aw + \frac{a^2}{2!}(w-1)^2 - \frac{a^3}{3!}(w-2)^3 + \ldots + (-)^r \frac{a^r}{r!}(w-r+1)^r + \ldots, \tag{9}$$

the series continuing so long as $w - r + 1 > 0$.

Equation (8) can, with a change of units, be put into the form of equation (1), based on Garwood's method.

It can be shown by differentiating (6) and (7), that

$$f'(w) = 0 \qquad (0 < w < 1), \tag{10}$$

and

$$f'(w) = -af(w-1) \quad (w > 1), \tag{11}$$

and hence that

$$f(w) = a\,\theta(w-1) - a\,e^{-q}\,\theta(w-2) \quad (w > 0, \neq 1). \tag{12}$$

In the ranges $(0, 1)$, $(1, 2)$, $(2, 3)$, ..., $f(w)$ is successively constant, a linear, quadratic, ... function of w; it is discontinuous at $w = 1$.

2·2·2. The moment-generating function (m.g.f.) $M_w(t)$ of the distribution of the delay is given by the Stieltjes integral $M_w(t) =$ expectation of e^{wt}

$$= -\int_0^\infty e^{wt}\,dP(w).$$

Using integration by parts and equations (5), (6) and (7), we obtain

$$M_w(t) = \frac{e^{-q-t}(q-t)}{q\,e^{-q} - t\,e^{-t}}. \tag{13}$$

Expansion of this function leads to the following expressions for the mean and variance of the distribution:

$$\mu_1'(w) = \frac{e^q - q - 1}{q}, \tag{14}$$

$$\mu_2(w) = \frac{e^{2q} - 2q\,e^q - 1}{q^2}. \tag{15}$$

In the original units, the mean delay will be

$$\mu_1'(T) = \frac{e^{NI} - NI - 1}{N}. \tag{16}$$

2·3. The third derivation of $P(w)$ is in one sense more straightforward; the solution is built up one step at a time.

2·3·1. Each vehicle passing X 'blocks' the road to pedestrians for a unit time before he reaches it; no pedestrian can start to cross in this interval. Let $k(s)$ $(s = 0, 1, 2, \ldots)$ be the probability that a random pedestrian has to wait for exactly s vehicles to pass before he can start to cross. This is the probability that the first s time-gaps between vehicles are all less than unity, but that the $s+1$th is greater than unity. This clearly has probability $(1 - e^{-q})^s e^{-q}$. Putting

$$\lambda = 1 - e^{-q}, \tag{17}$$

we obtain

$$k(s) = \lambda^s - \lambda^{s+1} \quad (s = 0, 1, \ldots). \tag{18}$$

Now, given that the pedestrian has to wait for exactly s vehicles to pass, the distribution of his delay, w_s, is the same as the distribution of the time before the sth vehicle arrives. This is the sum of s independent intervals, each of which has a distribution

$$\frac{q}{\lambda} e^{-qw}\,dw$$

between 0 and 1. The m.g.f. of this distribution is

$$M = \int_0^1 \frac{q}{\lambda} e^{-qw} e^{tw}\,dw$$

$$= \frac{q}{\lambda(t-q)}(e^{t-q} - 1). \tag{19}$$

Hence the m.g.f. of the distribution of w_s will be M^s. The m.g.f. of the delay, averaged over all values of s, is therefore

$$M_w(t) = \sum_{s=0}^{\infty} k(s)\, M^s$$

$$= \frac{e^{-q-t}(q-t)}{q e^{-q} - t e^{-t}}, \tag{20}$$

as before, substituting from (17), (18) and (19). The series converges if t is less than the root of

$$t e^{-t} = a,$$

other than $t = q$.

It should be possible to obtain the frequency function from this expression by application of the Laplace transformation, but a satisfactory method of doing this has not so far been found.

3. The size of groups of pedestrians

It is clear that pedestrians will cross the road in groups (actually in groups of arrival-groups, though at present we assume that pedestrians arrive singly), the groups tending to be larger when the vehicle flow or the pedestrian flow is larger. A proportion e^{-q} of pedestrians are not delayed, will not cross immediately behind vehicles, and will all cross in groups of size unity.

We define variates α, β, γ, δ and ϵ as follows:

α: the number of pedestrians waiting to cross when a random vehicle passes X; $A_0, A_1, A_2, \ldots$ are the probabilities that $\alpha = 0, 1, 2, \ldots$;

β: the number of pedestrians crossing immediately behind a random vehicle (probabilities $B_0, B_1, \ldots$);

γ: the size of a random group of pedestrians crossing (probabilities $C_1, C_2, \ldots$);

δ: the number of pedestrians waiting at a random time (probabilities $D_0, D_1, \ldots$);

ϵ: the size of group with which a random pedestrian crosses (probabilities $E_1, E_2, \ldots$).

Also let $p = nl$ be the average number of pedestrians arriving per unit time (on the changed time-scale).

We derive below the factorial moment-generating functions (f.m.g.f.'s) of α, β, γ, δ and ϵ, and hence their mean values. The complete probability distributions have not been evaluated, but the probabilities of 0 and 1 are given for each distribution.

3·1. Given that the road has been 'blocked' to pedestrians for a time w (at a general time this has the same distribution as the delay w), the probability that there will be r pedestrians waiting is

$$\frac{e^{-pw}(pw)^r}{r!} \quad (r = 0, 1, 2, \ldots).$$

The m.g.f. of r will be

$$M_{r|w}(t) = \sum_{r=0}^{\infty} \frac{e^{-pw}(pw)^r . e^{tr}}{r!}$$

$$= e^{pw(e^t-1)}. \tag{21}$$

Hence at a random time the m.g.f., $M_\delta(t)$, of δ, the number of pedestrians waiting, will be

$$-\int_0^\infty e^{pw(e^t-1)}\, dP(w).$$

This is clearly $M_w(\nu)$, where $\nu = p(e^t - 1)$. We therefore have

$$M_\delta(t) = \frac{e^{-q-\nu}(q-\nu)}{q e^{-q} - \nu e^{-\nu}}, \quad \text{from (13).} \tag{22}$$

Since

$$M_x(t) = F M_x(e^t - 1) = F M_x(\nu/p), \tag{23}$$

where $FM_x(t)$ is the f.m.g.f. of the distribution of x with argument t, we have

$$FM_\delta(t) = \frac{e^{-q-pt}(q-pt)}{q\,e^{-q}-pt\,e^{-pt}}. \tag{24}$$

3·2. To find the generating functions of the distribution of α, we note that when a vehicle is at X, the road has been 'blocked' for a time $1+w$, where w has m.g.f. given by (13); hence the f.m.g.f. of α is obtained by replacing w by $1+w$ in the above argument. Thus

$$M_\alpha(t) = M_{1+w}(\nu),$$

and so

$$FM_\alpha(t) = \frac{e^{-q}(q-pt)}{q\,e^{-q}-pt\,e^{-pt}}. \tag{25}$$

3·3. Next, to obtain the generating functions of the distribution of β, we note that $r\,(>0)$ pedestrians will cross behind a random vehicle if there are r waiting and if there is no vehicle in the following unit time. Hence

$$B_r = e^{-q}A_r \quad (r = 1, 2, \ldots), \tag{26}$$

and since

$$\sum_{r=0}^{\infty} B_r = \sum_{r=0}^{\infty} A_r = 1,$$

we must have

$$B_0 = e^{-q}A_0 + 1 - e^{-q}. \tag{27}$$

Therefore

$$M_\beta(t) = e^{-q}M_\alpha(t) + 1 - e^{-q},$$

and so

$$FM_\beta(t) = 1 - pt\,e^{-q}\frac{e^{-q}-e^{-pt}}{q\,e^{-q}-pt\,e^{-pt}}. \tag{28}$$

3·4. The generating functions of the distribution of γ may be obtained from those of β as follows: C_r will be proportional to B_r for $r = 2, 3, \ldots$, since all groups of size 2 or more must be behind a vehicle. Let, then,

$$C_1 = \kappa B_1 + \mu,$$
$$C_2 = \kappa B_2,$$
$$C_3 = \kappa B_3, \quad \text{and so on.}$$

We may find κ and μ as follows: We have

$$1 = \kappa(1-B_0) + \mu,$$

and

$$1 - e^{-q} = \kappa \sum_1^\infty rB_r \Big/ \left(\kappa \sum_1^\infty rB_r + \mu\right),$$

the second equation expressing the fact that a proportion $1-e^{-q}$ of pedestrians cross over immediately behind a vehicle.

Now

$$B_0 = (1-e^{-q}) + e^{-q}A_0$$

and

$$A_0 = -\int e^{-p(1+w)}\,dP(w)$$
$$= e^{-p}M_w(-p) = \frac{e^{-q}(q+p)}{q\,e^{-q}+p\,e^{p}}.$$

Thus

$$B_0 = 1 + p\,e^{-q}\frac{e^{-q}-e^{p}}{q\,e^{-q}+p\,e^{p}}.$$

Also

$$\sum_1^\infty rB_r = \mu_1'(\beta) = p(1-e^{-q})/q,$$

on expanding $M_\beta(t)$. The two equations for κ and μ can now be solved, giving

$$\frac{p\,e^{-q}}{q}\kappa = \mu = \frac{p\,e^{p}+q\,e^{-q}}{e^{p}\,(p+q)}. \tag{29}$$

We finally have, therefore, $M_\gamma(t) = \kappa(M_\beta(t)-B_0)+\mu\,e^{t}$,

leading to
$$FM_\gamma(t) = 1+\frac{pt-q}{pt\,e^{-pt}-q\,e^{-q}}\,\frac{t\,e^{-pt}}{e^{p}}\,\frac{q\,e^{-q}+p\,e^{p}}{q+p}. \tag{30}$$

3·5. To obtain the generating functions of the distribution of ϵ, we note that

$$E_r = rC_r \Big/ \sum_1^\infty iC_i \quad (r = 1, 2, \ldots). \tag{31}$$

Hence
$$M_\epsilon(t) = \frac{1}{\mu_1'(\gamma)}\sum_1^\infty C_r r\,e^{rt}$$
$$= \frac{1}{\mu_1'(\gamma)}\,\frac{dM_\gamma(t)}{dt},$$

which on evaluation leads to

$$FM_\epsilon(t) = \frac{(1+t)\,e^{-q-pt}}{(pt\,e^{-pt}-q\,e^{-q})^2}\,[p^2t^2\,e^{-pt}+q\,e^{-q}(-2pt+q+p^2t^2-pqt)]. \tag{32}$$

Table 1. *Mean values of* α, β, γ, δ, ϵ, *and probabilities of* 0 *and* 1

x	Mean, $\bar{x}$	Probability ($x=0$)	Probability ($x=1$)
α	$\frac{p(e^q-1)}{q}$	$\frac{e^{-q}(q+p)}{q\,e^{-q}+p\,e^{p}}$	$\frac{p\,e^{-q}\,[-q\,e^{-q}+e^{p}(p^2+pq+q)]}{[q\,e^{-q}+p\,e^{p}]^2}$
β	$\frac{p(1-e^{-q})}{q}$	$1-p\,e^{-q}\,\frac{e^{p}-e^{-q}}{q\,e^{-q}+p\,e^{p}}$	$\frac{p\,e^{-2q}\,[-q\,e^{-q}+e^{p}(p^2+pq+q)]}{[q\,e^{-q}+p\,e^{p}]^2}$
γ	$\frac{qe^{-q}+pe^{p}}{e^{p-q}(q+p)}$	0	$\frac{p}{p+q}+\frac{q\,e^{-q}(1+p)}{q\,e^{-q}+p\,e^{p}}$
δ	$\frac{p}{q}(e^q-q-1)$	$\frac{e^{p-q}(q+p)}{q\,e^{-q}+p\,e^{p}}$	$\frac{pq\,e^{p-q}[e^q-(p+q+1)\,e^{-q}]}{[q\,e^{-q}+p\,e^{p}]^2}$
ϵ	$1+\frac{2p}{q}(e^q-q-1)$	0	$\frac{e^{p-q}[p^2\,e^{p}+q\,e^{-q}(2p+q+p^2+pq)]}{[q\,e^{-q}+p\,e^{p}]^2}$

3·6. We can write down the probability-generating functions (p.g.f.'s) of α, β, γ, δ and ϵ quite simply in terms of the f.m.g.f.'s. For any distribution in which the variable takes only positive integral values we have the identity

$$\text{p.g.f. of } x = P_x(t) = p_0+tp_1+t^2p_2+\ldots$$
$$= FM_x(t-1),$$

where p_r = probability that $x = r$.

The first few terms of the moment- and probability-generating functions have been expanded and Table 1 gives the expressions obtained. Figs. 1 (*a*)–(*d*) show how the means vary with p and q for the ranges of pedestrian and traffic flow $0 \leqslant p \leqslant 2, 0 \leqslant q \leqslant 2$.

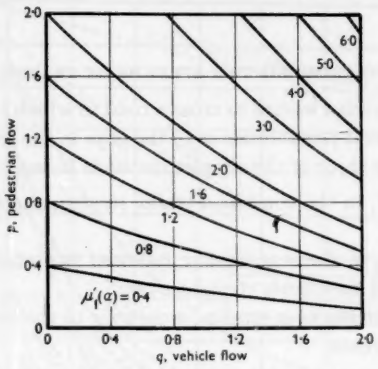


Fig. 1 *a*. Mean value of α, the number of pedestrians waiting to cross when a random vehicle passes.

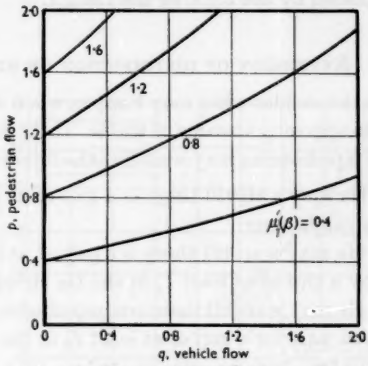


Fig. 1 *b*. Mean value of β, the number of pedestrians crossing behind a random vehicle.

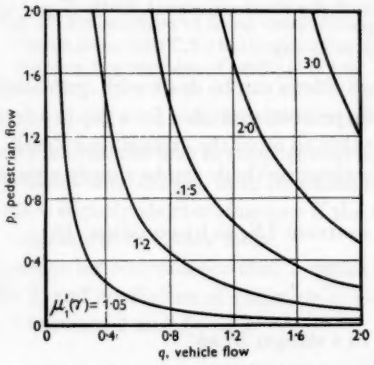


Fig. 1 *c*. Mean value of γ, the size of a random group of pedestrians crossing.

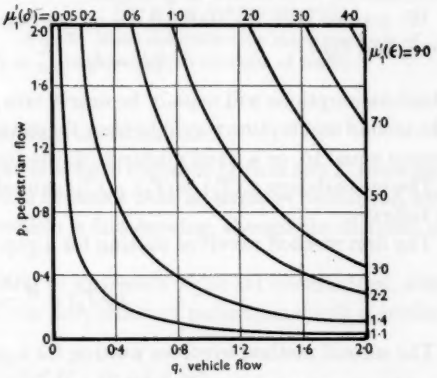


Fig. 1 *d*. Mean values of (i) δ, the number of pedestrians waiting at a random time, (ii) ϵ, the size of group in which a random pedestrian crosses.

3·7. The distributions just obtained for $\alpha, \beta, \gamma, \delta$ and ϵ can be extended quite simply to the case when pedestrians arrive in independent groups, provided the distribution of the size of these is known.

Let g be the number of pedestrians in a random arrival-group, and let $M_g(t)$ and $K_g(t)$ be the m.g.f. and cumulant-generating function (c.g.f.) of its distribution.

Consider, for example, the distribution of α. Suppose that $\alpha = r$. Then the m.g.f. of the number of individuals in this group will be $[M_g(t)]^r$, or $e^{rK_g(t)}$. If we now average this m.g.f. over the distribution of α, we obtain the m.g.f. of the number of individuals waiting when a random vehicle passes as

$$A_0 e^0 + A_1 e^{K_g(t)} + A_2 e^{2K_g(t)} + \ldots = M_\alpha(K_g(t)),$$

i.e. to obtain the m.g.f. of the number of individuals, we replace t by $K_g(t)$ in the m.g.f. of the number of arrival groups.

The same rule can easily be seen to apply with β, γ, δ or ϵ in place of α, and the m.g.f. may be replaced by the c.g.f. or the f.m.g.f.

4. Extension of distribution of delay to crossing two streams of traffic

We next consider what may happen when a pedestrian wishes to cross a road in which there are two opposing streams of traffic. There are various possibilities as to the gaps in the traffic which a pedestrian may wait for; the following are three of the simpler methods he may use:

(1) He may wait till there is a gap of at least I_1 in the combined traffic stream, and then cross straight over.

(2) He may wait till there is a gap of at least I_2 in the near stream, followed immediately after by a gap of at least I_2' in the far stream, and then cross straight over.

(3) He may wait till there is a gap of at least I_3 in the near stream, cross over to the centre and then wait for a gap of at least I_3' in the far stream.

The third method will normally be used when there is a refuge in the centre of the road. When there is not, it is not obvious whether the first or second method is a better representation of what occurs in practice.

We assume that the traffic flows on the two halves of the road are equal, both N per unit time, and that

$$\tfrac{1}{2}I_1 = I_2 = I_2' = I_3 = I_3', = I, \text{ say.}$$

These assumptions will usually be nearly true, though others can be dealt with quite easily. The second assumption would be true, for example, if a pedestrian waited for a gap in a traffic stream equal to, or a fixed multiple of, the time he takes to cross the stream completely.

The mean delays $\mu_1'(T_1), \mu_1'(T_2), \mu_1'(T_3)$ caused by the three methods can be simply obtained as follows:

The first method involves waiting for a gap $2I$ in a stream $2N$, so by equation (16)

$$\mu_1'(T_1) = \frac{(e^{4NI} - 4NI - 1)}{2N}. \tag{33}$$

The second method involves waiting for a gap $2I$ in a stream N, so

$$\mu_1'(T_2) = \frac{(e^{2NI} - 2NI - 1)}{N}. \tag{34}$$

The third method involves waiting twice for gaps I in streams N, so

$$\mu_1'(T_3) = \frac{2(e^{NI} - NI - 1)}{N}. \tag{35}$$

The corresponding m.g.f.'s may be written down if required.

Fig. 2 shows the probability of being delayed for longer than T using each of the three methods, for various traffic flows and $I = 5$ sec. Fig. 3 shows the mean delays caused by each of the three methods, for various traffic flows and $I = 3$ and 5 sec. These curves show clearly that installation of refuges decreases the delay (since method (3) can be used), and also that the mean delay increases much more than proportionately with vehicle flow. However, when the flow is very heavy, the assumption of randomness in the arrival of vehicles breaks down, probably to the advantage of the pedestrian.

The sizes of groups of pedestrians, corresponding to α, β, γ, δ and ϵ already considered, have not been investigated in the case of two streams of traffic, though some of the results for a single stream apply directly.

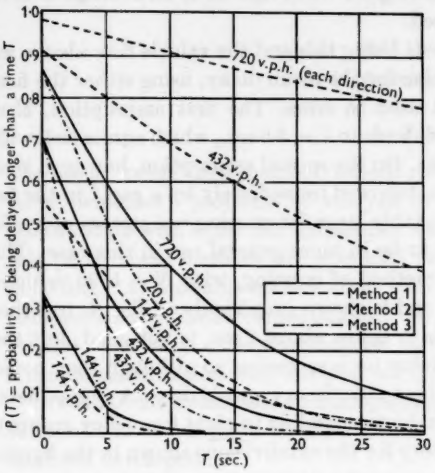


Fig. 2. Probabilities of being delayed longer than a given time T, for the three methods of crossing two streams of traffic ($I = 5$ sec.).

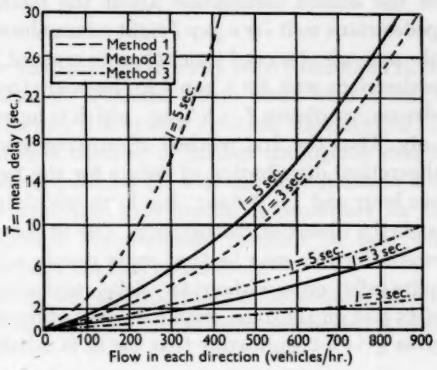


Fig. 3. Mean delays for the three methods of crossing two streams of traffic.

5. THE EFFECT OF A VARIABLE CRITICAL GAP I

It is, of course, not true to suppose that all pedestrians have the same critical gap I; some can walk faster than others, and, in addition, there is evidence that an average pedestrian will accept a slightly shorter time gap if the next vehicle is fast-moving, though the distance of the vehicle will be rather more.

Assume for convenience that I varies according to Pearson's Type III distribution, with mean I_0 and coefficient of variation v, i.e. that the proportion of pedestrians with a critical gap I between I and $I+dI$ is

$$\frac{1}{(I_0 v^2)^{1/v^2}\,\Gamma(1/v^2)}\, I^{1/v^2-1}\, e^{-I/I_0 v^2}\, dI \quad (0 < I < \infty).$$

The mean delay to pedestrians crossing one stream of traffic will now be

$$\bar{T} = \int_0^\infty \frac{e^{NI} - NI - 1}{N} \frac{1}{(I_0 v^2)^{1/v^2}\,\Gamma(1/v^2)}\, I^{1/v^2-1}\, e^{-I/I_0 v^2}\, dI.$$

This may easily be evaluated as

$$\bar{T} = \frac{(1 - NI_0 v^2)^{-1/v^2} - NI_0 - 1}{N}. \tag{36}$$

As $v \to 0$, this expression tends to the value given by (16).

As an example, if we take $N = 360$ vehicles per hour, $I_0 = 5$ sec., $v = 0{\cdot}0$, $0{\cdot}2$ and $0{\cdot}4$, the mean delays are

$$v = 0{\cdot}0 : \bar{T} = 1{\cdot}49 \text{ sec.}$$
$$0{\cdot}2 : \bar{T} = 1{\cdot}57 \text{ sec.}$$
$$0{\cdot}4 : \bar{T} = 1{\cdot}84 \text{ sec.}$$

6. Comparison of theory with some observations

Some measurements of pedestrian delays have been made at a pedestrian crossing without a refuge. The traffic remained fairly constant during the observations at an average level of 1090 vehicles per hour, both directions combined.

The mean delay to 325 pedestrians was 9·2 sec. Using this and the vehicle flow above, we can calculate the critical gap I from the formulae for the mean delay, using either the first or the second assumption about the method used to cross. The first assumption, that pedestrians wait for a gap I right across the road, leads to $I = 5{\cdot}6$ sec., which agrees well with the average observed time taken to cross of 7 sec. On the second assumption, however, that pedestrians wait for a gap I in the near stream, followed immediately by a gap I in the far stream, we obtain $I = 8{\cdot}7$ sec., which is unreasonably large for crossing one stream of traffic only. Thus the first method of crossing seems to be in more general use in this case. The theoretical distribution of delays for the first method of crossing, with $N = 1090$ vehicles per hour and $I = 5{\cdot}6$ sec., has been calculated and is shown graphically in Fig. 4, together with the observed distribution. The fit for delays below about 2 sec. is not good, but the reason for this may be that some people will delay for a second or so although they could quite safely cross without any delay, and according to our theory, would do so. A χ^2 goodness-of-fit test on the data, with 6 degrees of freedom (corresponding to eight frequency groups), gives $\chi^2 = 6{\cdot}2$, indicating that the fit is satisfactory for the subdivisions shown in the figure.

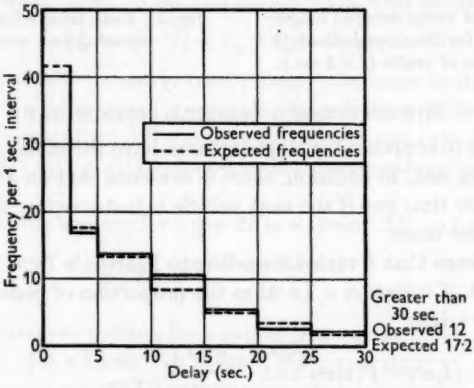


Fig. 4. Comparison of observed and theoretical distributions of delay.

This paper was prepared as part of the programme of the Road Research Laboratory and is published by permission of the Director of Road Research.

The author's sincere thanks are due to Dr F. Garwood for introducing him to the subject and for much help and advice, to Mr D. G. Kendall for several suggestions, and to the Road Users' Section of the Laboratory (under Mr R. L. Moore) for providing the data used in §6.

REFERENCES

Adams, W. F. (1936). *J. Instn Civ. Engrs*, Nov. 1936, p. 121.
Garwood, F. (1940). *J.R. Statist. Soc., Suppl.*, **7**, 1, 65.
Raff, M. S. (1951). *J. Amer. Statist. Assoc.* **46**, 253.

FURTHER APPLICATIONS OF RANGE TO THE ANALYSIS OF VARIANCE

By H. A. DAVID
University College, London

1. Introduction

The use of 'range' as a short-cut substitute for variance has recently been considered in the more elementary problems of analysis of variance. Patnaik (1950) dealt with a simple classification with equal cell frequencies and Hartley (1950) with a double classification. In this paper these methods are extended to cover a number of further orthogonal designs including: (*a*) a double classification with cell replication; (*b*) a double classification with two-factor 'treatment combinations'; (*c*) the split-plot experiment. A procedure for the general three-way cross-classification is indicated. Approximate methods to deal with a simple classification with unequal cell frequencies are also considered.

In each of the main cases an example is given as well as a summary of procedure, which it is hoped will be adequate for the reader interested only in the practical applications of the methods developed.

2. Double classification with cell replication

The essential difference between range methods and the orthodox analysis of variance consists, it will be recalled, in the replacement of mean-square estimators of variance by statistics involving ranges of observations in subgroups or of group means. We first describe the procedure in terms of an example.

(2·1). *Example* 1

Tables 1*a* and 1*b* form a convenient summary of an experiment on Portland cement (Davies, 1947, p. 91). Three 'gaugers' prepared twelve samples of Portland cement each, four of which were tested for compressive strength by each of three 'breakers'. Table 1*a* gives the totals, and Table 1*b* the ranges in lb./sq.in. of four test results for the nine breaker × gauger combinations.

Table 1*a*. *Totals of four replicates*

Gauger \ Breaker	1	2	3	Total
1	136	− 4	− 74	58
2	18	138	−194	− 38
3	270	166	− 33	403
Total	424	300	−301	423

Origin at 5000 lb./sq.in.

Table 1*b*. *Ranges of four replicates*

Gauger \ Breaker	1	2	3
1	104	186	116
2	116	132	62
3	80	96	114

Estimation of residual error. From Table 1*b* the mean range is given by

$$\overline{w} = (104 + \ldots + 114)/9 = 111{\cdot}8,$$

and the estimate s_w of the error standard deviation is at once obtained as $111{\cdot}8/2{\cdot}08 = 53{\cdot}8$, the divisor, $c = 2{\cdot}08$, as well as the equivalent degrees of freedom, $\nu = 24{\cdot}9$, of the estimate being obtained from Table I as described in the Appendix.* (From the within-groups sum of squares Davies finds $s = 51{\cdot}4$ with 27 degrees of freedom.)

Estimation of interaction. Table 1*a* is clearly an ordinary double classification of the group totals. Following Hartley (1950), we set out in Table 1*c* the 'breaker residuals' on which the estimate of the interaction term† of such a design is based (for example, the first entry, -5, is given to the nearest integer by $136 - 424/3$). The figures in each column sum to zero except for rounding-off errors, which provides a useful check. If there is no real interaction we can, as usual, obtain an estimate s'_w of the error S.D. from the interaction term. Remembering that Table 1*c* gives totals, this will be $s'_w = 354 \times \sqrt{4}/(12 \times 1{\cdot}48) = 39{\cdot}9$; the corrective factor $c' = 1{\cdot}48$ and the equivalent degrees of freedom, $\nu' = 3{\cdot}7$, are read off this time from Table II.

Table 1*c*. *Breaker residuals (totals)*

Breaker / Gauger	1	2	3	Range
1	− 5	−104	26	130
2	−123	38	−94	161
3	129	66	67	63
Total				354

Test for interaction and main factors. s'^2_w/s^2_w with $s_w = 53{\cdot}8$ and $s'_w = 39{\cdot}9$ may now be referred to tables of the F-ratio with degrees of freedom $\nu' = 3{\cdot}7$, $\nu = 24{\cdot}9$. The interaction is obviously not significant.

To test for the main gauger and breaker factors we form 'Studentized' ranges q_A, q_B. Thus for gaugers,

$$q_A = \{403 - (-38)\}/(\sqrt{(12)} \times 53{\cdot}8) = 2{\cdot}37,$$

and for breakers

$$q_B = \{424 - (-301)\}/(\sqrt{(12)} \times 53{\cdot}8) = 3{\cdot}89.$$

On consulting tables of percentage points of q (Pearson & Hartley, 1943; Hartley, 1950) with $\nu = 24{\cdot}9$ and sample size 3, only q_B will be found significant (5 % point, 3·53, 1 % point, 4·54).

These three tests correspond to those made by Davies and give in each case the same result. It may be noted that the labour involved in using the range method is, in contrast to the standard procedure, little affected by an increase in the number of replications.

* It will be remembered that the corrective factor c is used to give an estimate, the expected value of whose square is σ^2. The commonly used corrective factor d_n leads to an unbiased estimate of σ. As the number of observations n increases, $d_n \to c$.

† In the ordinary double classification this term provides the 'error'.

(2·2). *Theoretical basis*

Just as in the orthodox analysis of a double classification with cell replication, it is necessary to distinguish between the 'systematic' and 'random' probability models. We discuss first the systematic case, the analysis of which has been illustrated by the preceding example.

Systematic model. This set-up may be written as

$$x_{ijt} = K + A_i + B_j + I_{ij} + z_{ijt} \quad (i = 1 \dots l;\ j = 1 \dots m;\ t = 1 \dots n), \tag{1}$$

where capital letters denote systematic effects which are subject to the constraints

$$\sum_{i=1}^{l} A_i = 0 = \sum_{j=1}^{m} B_j \quad \text{and} \quad \sum_{i=1}^{l} I_{ij} = 0 = \sum_{j=1}^{m} I_{ij}.$$

The z_{ijt} are mutually independent normal variates with zero means and variance σ^2. We then have the fundamental equation

$$w_n(i,j) \equiv \text{range}\,(x_{ijt}) = \text{range}\,(z_{ijt}), \quad \text{for fixed } i,j \text{ and } t = 1 \dots n, \tag{2}$$

so that the statistic $w_n(i,j)$ can be computed as the range of observations x_{ijt} and its distribution regarded as that of the range of the random variables z_{ijt}. Patnaik (1950) has approximated to the distribution of the mean $\overline{w}$ of a set of independent ranges by equating the expectation M and variance V of $\overline{w}/\sigma$ to those of $c\chi/\sqrt{\nu}$, where χ has ν degrees of freedom. Table I, giving the constants c, ν, is an extension of Patnaik's Table 1, and is to be entered here with sample size n, number of samples $k = lm$ to give $s_w = \overline{w}/c$, the estimator of σ.

From (1), $\bar{x}_{ij.} = K + A_i + B_j + I_{ij} + \bar{z}_{ij.}$. Corresponding to (2), we reach the equation

$$w'_l(j) \equiv \text{range}\,(\bar{x}_{ij.} - \bar{x}_{i..}) = \text{range}\,(I_{ij} + \bar{z}_{ij.} - \bar{z}_{i..}), \quad \text{for fixed } j \text{ and } i = 1 \dots l. \tag{3}$$

The m ranges w'_l are not independent. However, for the null-hypothesis

$$I_{ij} = 0 \quad (i = 1 \dots l;\ j = 1 \dots m)$$

Hartley (1950) obtained the mean and variance of $\overline{w}'$ and used these to fit a similar χ-approximation to the mean of the correlated ranges. His table of the resulting ν' and c' is extended below in Table II. It is clear that on the null-hypothesis $s'_w = \overline{w}'\sqrt{(n)}/c'$ is an estimator of σ irrespective of whether A_i and/or B_j are present. On the alternative hypothesis (I_{ij} not all zero) s'_w will be seen from (3) to depend on the interaction terms I_{ij} and the z_{ijt} only. It may be shown that s_w and s'_w are independent (see §4), so that on the null-hypothesis s'^2_w/s^2_w is distributed approximately as F with degrees of freedom ν', ν. Finally, since $\bar{x}_{i..} = K + A_i + \bar{z}_{i..}$, the main A effects may be tested by the usual 'Studentized' range criterion; thus

$$q_A = \sqrt{(mn)} \times \text{range}\,(\bar{x}_{i..})/s_w, \quad q_B = \sqrt{(ln)} \times \text{range}\,(\bar{x}_{.j.})/s_w.$$

*Random model.** For this case we may write

$$x_{ijt} = K + a_i + b_j + u_{ij} + z_{ijt} \quad (i = 1 \dots l;\ j = 1 \dots m;\ t = 1 \dots n), \tag{4}$$

where the $a_i, b_j, u_{ij}, z_{ijt}$ are all mutually independent normal variates with zero means and respective variances $\sigma_a^2, \sigma_b^2, \sigma'^2, \sigma^2$. From a slight modification of the above arguments it follows that, as in the ordinary analysis of variance, we proceed in the same way as for the systematic case, except that main factors must now be tested against s'_w so that

$$q_A = \sqrt{(mn)}\,\text{range}\,(\bar{x}_{i..})/s'_w.$$

* For the null-tests, the essential feature of the random model is the random interaction term u_{ij}. The main effects may be systematic, or random as in (4).

(2·3). *Summary of procedure*

Results may conveniently be set out in Table 2.

Table 2. *Summary table for double classification with cell replication*

Source	Function calculated	Sample size	Degrees of freedom	Results for Example 1		
				Numerical value of function	Sample size	Degrees of freedom
(i) Treatment A	$\sqrt{(mn)}$ range $(\bar{x}_{i..})$	l	—	$\sqrt{(12)} \times 441/12$	3	—
(ii) Treatment B	$\sqrt{(ln)}$ range $(\bar{x}_{.j.})$	m	—	$\sqrt{(12)} \times 725/12$	3	—
(iii) Interaction	$s'_w = w'\sqrt{(n)}/c'$	—	ν' (Table II)	$(354/12) \times \sqrt{4/1{\cdot}48}$	—	3·7
(iv) Residual	$s_w = \bar{w}/c$	—	ν (Table I)	111·8/2·08	—	24·9

Tests

For treatment A:

Random model: refer (i)/(iii) to tables of q with sample size l and D.F $= \nu'$,

Systematic model: refer (i)/(iv) to tables of q with sample size l and D.F. $= \nu$.

For interaction refer, on either model, $(\text{iii})^2/(\text{iv})^2$ to tables of F with degrees of freedom ν, ν'.

(2·4). *Simple hierarchical classification**

Another set-up which may readily be handled by the methods so far developed is the simple hierarchical classification represented by the model

$$x_{ijt} = K + A_i + B_{ij} + z_{ijt} \quad (i = 1 \dots l;\ j = 1 \dots m;\ t = 1 \dots n). \tag{5}$$

3. Double classification, 'treatments' × 'blocks', with two-factor 'treatment combinations'

For definiteness we consider the theoretical model in the systematic form

$$x_{ijt} = K + A_i + B_j + I_{ij} + C_t + z_{ijt}, \tag{6}$$

where the suffices (i, j) denote the treatment combinations which may be broken up into two treatments and their interaction and t is the block index.

Since from (6)
$$\bar{x}_{ij.} = K + A_i + B_j + I_{ij} + \bar{z}_{ij.}, \tag{7}$$

it follows that the residual error may be estimated in the same way as the interaction (blocks × treatment combinations) in an $lm \times n$ ordinary double classification. The I_{ij} interaction term is estimated as in §2.

(3·1). *Example* 2

The full procedure is set out in Table 3*a* (Paterson, 1939, p. 206).

The block means are: $\bar{x}_{..1} = 217$, $\bar{x}_{..2} = 200$, $\bar{x}_{..3} = 201$, $\bar{x}_{..4} = 215$.

We now form the block residuals $x_{ijt} - \bar{x}_{..t}$ $(i = 1 \dots 3;\ j = 1 \dots 4)$ as in Table 3*b*. Hence $\bar{w} = 735/12 = 61{\cdot}25$. From Table II, $c = 1{\cdot}98$, $\nu = 24{\cdot}4 + 5{\cdot}4 = 29{\cdot}8$, so that $s_w = 30{\cdot}9$. (Paterson obtains $s = 32{\cdot}2$ with 33 D.F.) The interaction may be computed most conveniently

* For an example see Davies (1947, p. 81).

(as in § 2) from the totals $4(\bar{x}_{ij.} - \bar{x}_{.j.})$, resulting in the ratio $1229 \times \sqrt{4}/(30{\cdot}9 \times 12 \times 1{\cdot}76) = 3{\cdot}77$, whose square when referred to the F-tables with degrees of freedom 5·6, 29·8 is highly significant. We also find

$$q_A = 2260/\sqrt{6 \times 30{\cdot}9} = 18{\cdot}3, \quad q_B = 21{\cdot}1 \quad \text{and} \quad q_C = 1{\cdot}85,$$

so that all factors except blocks are highly significant, a result identical with that reached by Paterson.

Table 3*a*. *Yields (in lb. of dry matter per 1/57 acre) of three varieties of grass with four cutting rotations, repeated in four blocks*

Grass ...	1				2				3				Block total	Rotation total
Rotation ...	1	2	3	4	1	2	3	4	1	2	3	4		
Block 1	96	187	222	109	146	252	246	277	115	298	220	430	2598	A = 1397
2	70	163	125	97	133	181	263	293	143	220	341	371	2400	B = 2582
3	77	143	134	133	154	224	194	260	117	234	258	484	2412	C = 2663
4	80	179	173	113	146	248	190	325	120	253	297	460	2584	D = 3352
Totals	323	672	654	452	579	905	893	1155	495	1005	1116	1745	9994	
	2101				3532				4361					

Table 3*b*. *Block residuals and their treatment ranges*

Grass ...	1				2				3			
Rotation ...	1	2	3	4	1	2	3	4	1	2	3	4
Block 1	−121	−30	5	−108	−71	35	29	60	−102	81	3	213
2	−130	−37	−75	−103	−67	−19	63	93	− 57	20	141	171
3	−124	−58	−67	− 68	−47	23	− 7	59	− 84	33	57	283
4	−135	−36	−42	−102	−69	33	−25	110	− 95	38	82	245
Range	14	28	80	40	24	54	88	51	45	61	138	112

(3·2). *Summary of procedure*

Results may be tabulated as in the preceding paragraph (see Table 4).

Tests

For treatment A:

Interaction random: refer (i)/(iv) to tables of q with sample size l and D.F. $= \nu_2'$.

Interaction systematic: refer (i)/(v) to tables of q with sample size l and D.F. $= \nu_1'$.

For blocks C refer (iii)/(v) to tables of q with sample size n and D.F. $= \nu_1'$ in either case.

For interaction refer $(\text{iv})^2/(\text{v})^2$ to tables of F with degrees of freedom ν_2', ν_1'.

Table 4. *Summary table for double classifications with two-factor treatment combinations*

Source	Function calculated	Sample size	Degrees of freedom	Results for Example 1		
				Numerical value of function	Sample size	Degrees of freedom
(i) Treatment A	$\sqrt{(mn)}$ range $(\bar{x}_{i..})$	l	—	565	3	—
(ii) Treatment B	$\sqrt{(ln)}$ range $(\bar{x}_{.j.})$	m	—	652	4	—
(iii) Blocks C	$\sqrt{(lm)}$ range $(\bar{x}_{..t})$	n	—	57·2	4	—
(iv) Interaction A × B	$s'_{w_2} = \bar{w}'_2\sqrt{n}/c'_2$	—	ν'_2 (Table II)	116	—	5·6
(v) Residual	$s'_{w_1} = \bar{w}'_1/c'_1$	—	ν'_1 (Table II)	30·9	—	29·8

4. The split-plot experiment*

The theoretical model appropriate to this design contains at least two random terms and may be written as

$$x_{ijt} = K + A_i + B_j + u_{ij} + C_t + I_{it} + z_{ijt}, \tag{8}$$

where the symbols have their previous meaning so that u_{ij}, z_{ijt} give rise to two independent error terms, here usually called error (1) and error (2), respectively. The only new feature to which this set-up gives rise is a slight modification of the way in which the residual (error (2)) is estimated.

Estimation of σ. From (8) we have

$$x_{ijt} - \bar{x}_{ij.} = C_t + I_{it} + z_{ijt} - \bar{z}_{ij.},$$

and consequently

$$w_m(i,t) \equiv \text{range}\,(x_{ijt} - \bar{x}_{ij.}) = \text{range}\,(z_{ijt} - \bar{z}_{ij.}), \text{ for fixed } i, t \text{ and } j = 1 \ldots m.$$

Thus we obtain ln correlated ranges $w_m(i,t)$. The estimate of σ will again be $\bar{w}/c$ with ν degrees of freedom, ν and c being in this case obtainable from Table III (see Example 3) whose derivation is explained in the Appendix.

(4·1). *Theoretical basis*

In §2 we mentioned Patnaik's (1950) approximation to the distribution of the mean of independent ranges by fitting a χ-distribution adjusted to have the correct mean M and variance V. If the same approximation is made for the mean of correlated ranges the resulting ν and c may be determined quite generally from his equations

$$\frac{V}{M^2} = \frac{1}{2\nu} + \frac{1}{8\nu^2} - \frac{1}{16\nu^3} + \ldots, \tag{9}$$

and

$$c = M\left(1 + \frac{1}{4\nu} + \frac{1}{32\nu^2} - \frac{5}{128\nu^3}\cdots\right), \tag{10}$$

provided that for any given design we find the appropriate expressions for M and V.

For the split-plot design the correlation between $z_{ijt} - \bar{z}_{ij.}$ and $z_{i'jt} - \bar{z}_{i'j.}$ is zero if $i \neq i'$ and is $-1/(n-1)$ if $i = i'$, so that the ln ranges fall into l independent groups of n. Within each

* For a description and application of this design see, for example, Snedecor (1946, p. 309).

group any two ranges are correlated with correlation coefficient ρ_w tabled by Hartley (1950, Table 4). Here ρ_w is a function of $\rho = -1/(n-1)$ and the sample size m. Therefore

$$M \equiv \mathscr{E}(\overline{w}) = d_m \sigma (1-n^{-1})^{\frac{1}{2}}, \tag{11}$$

$$V \equiv \text{var}(\overline{w}) = \text{var}\left(\sum_i \sum_t w_m(i,t)/ln\right)$$

$$= (1+\overline{n-1}\rho_w) V_m \sigma^2 (1-n^{-1})/(ln), \tag{12}$$

and

$$\frac{V}{M^2} = \frac{1}{ln}\frac{V_m}{d_m^2}[1+(n-1)\rho_w(n,m)], \tag{13}$$

where d_m is the expectation and V_m the variance of the range of m independent unit normal variates. From (9), (10), (11) and (13) Table III can now be constructed.

Before giving an example we shall prove that each of the test criteria employed has independent numerator and denominator, even when the null-hypothesis does not hold (i.e. some or all of the effects are real). The method of proof is quite general (compare Hartley, 1950). From the ordinary analysis of variance, it follows that for the present set-up

$$\begin{aligned}\Sigma x_{ijt}^2 = {} & \Sigma(z_{ijt}-\bar{z}_{i.t}-\bar{z}_{ij.}+\bar{z}_{i..})^2+\Sigma(I_{it}+\bar{z}_{i.t}-\bar{z}_{i..}-\bar{z}_{..t}+\bar{z})^2 \\ & +\Sigma(C_t+\bar{z}_{..t}-\bar{z})^2+\Sigma(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y})^2 \\ & +\Sigma(B_j+\bar{y}_{.j}-\bar{y})^2+\Sigma(A_i+\bar{y}_{i.}-\bar{y})^2+N\bar{x}^2,\end{aligned}$$

where Σ denotes summation over all observations and $y_{ij} = u_{ij}+\bar{x}_{ij.}$. Reducing the number of variates in each Σ-term to its degrees of freedom and introducing

$$v_{ijt} = z_{ijt}-\bar{z}_{ij.}-\bar{z}_{i.t}+\bar{z}_{i..} \quad (i = 1 \dots l;\ j = 2 \dots m;\ t = 2 \dots n),$$

$$Z_{it} = I_{it}+\bar{z}_{i.t}-\bar{z}_{i..}-\bar{z}_{..t}+\bar{z} \quad (i = 2 \dots l;\ t = 2 \dots n) \text{ etc.},$$

the first two terms become

$$\sum_{i=1}^{l}\{\sum_j \sum_t v_{ijt}^2+\sum_j(\sum_t v_{ijt})^2+\sum_t(\sum_j v_{ijt})^2+(\sum_{t,j} v_{ijt})^2\}+\sum_{i=2}^{l}\sum_t Z_{it}^2+\sum_i(\sum_t Z_{it})^2+\sum_t(\sum_i Z_{it})^2+(\sum_{i,t} Z_{it})^2,$$

and it is clear that the joint distribution of the v_{ijt} is independent of that of the Z_{it}, etc. But the ranges of $z_{ijt}-\bar{z}_{ij.}$ may be defined in terms of the new variables v_{ijt}, so that they are distributed independently of the ranges of the $I_{it}+\bar{z}_{i.t}-\bar{z}_{i..}$, which can be defined in terms of the Z_{it}, etc. The general result follows.

(4·2). *Example* 3

Table 5 *a* is a section of Snedecor's table (1946, p. 310) of the yields in tons/acre of three varieties, planted in six blocks and subject to four subtreatments. This table, giving the yields of one of the varieties is simply a 4×6 double classification. Subtracting the subtreatment mean from each entry we obtain Table 5*c*, in which the ranges of the residuals in each block have been formed. In a similar way Table 5*d* is derived from the table of means (5 *b*) and enables us to estimate the variety-treatment interaction.

Estimation of residual error. We have $l = 3$, $m = 6$, $n = 4$. $\overline{w}$ is the mean of eighteen ranges of which the first six are listed in Table 5*c*. Using all varieties we find $\overline{w} = 0{\cdot}310$ and from Table III that $c = 1{\cdot}90$, giving $s_w = \overline{w}/c = 0{\cdot}163$. ν is obtained by multiplying $\nu' = 13{\cdot}4$ by three,* the number of varieties, so that $\nu = 40{\cdot}2$. (Snedecor finds $s = 0{\cdot}161$ with 45 degrees of freedom.)

* For a justification of this approximate procedure, see the Appendix (Table III).

Table 5*a*. *Yields of variety*

Sub-treatment	Block						Mean
	1	2	3	4	5	6	
1	2·17	1·88	1·62	2·34	1·58	1·66	1·88
2	1·58	1·26	1·22	1·59	1·25	0·94	1·31
3	2·29	1·60	1·67	1·91	1·39	1·12	1·66
4	2·23	2·01	1·82	2·10	1·66	1·10	1·82

Table 5*b*. *Variety-subtreatment means*

Sub-treatment	Variety (main treatment)			Mean
	1	2	3	
1	1·88	1·76	1·70	1·78
2	1·31	1·30	1·41	1·34
3	1·66	1·58	1·48	1·57
4	1·82	1·64	1·61	1·69
Mean	1·67	1·57	1·55	

Table 5*c*. *Subtreatment residuals of Table 5a and their block ranges*

Sub-treatment	Block					
	1	2	3	4	5	6
1	0·29	0·00	−0·26	0·46	−0·30	−0·22
2	0·27	−0·05	−0·09	0·28	−0·06	−0·37
3	0·63	−0·06	0·01	0·25	−0·27	−0·54
4	0·41	0·19	0·00	0·28	−0·16	−0·72
Range	0·36	0·25	0·27	0·21	0·24	0·50

Table 5*d*. *Table for the estimation of the variety-treatment interaction*

Sub-treatment	Variety			Range
	1	2	3	
1	0·21	0·19	0·15	0·06
2	−0·36	−0·27	−0·14	0·22
3	−0·01	0·01	−0·07	0·08
4	0·15	0·07	0·06	0·09

Estimation of variety-subtreatment interaction. From Table 5*d* we have $4\overline{w}'_{\mathrm{I}} = 0{\cdot}45$. Table II gives $c'_{\mathrm{I}} = 1{\cdot}54$, $\nu'_{\mathrm{I}} = 5{\cdot}4$. Hence $s'_{\mathrm{I}} = \overline{w}'_{\mathrm{I}}\sqrt{(6)}/c'_{\mathrm{I}} = 0{\cdot}179$, so that $F_{5\cdot4,\,40\cdot2} = 1{\cdot}24$ (corresponding to Snedecor's $F_{6,45} = 1{\cdot}25$). The remaining procedure will be clear from the summary table (Table 6).

(4·3). *Summary of procedure*

Table 6. *Summary table for the split-plot design*

Source	Function calculated	Sample size	Degrees of freedom
(i) Main treatment A	$\sqrt{(mn)}$ range $(\bar{x}_{i..})$	l	—
(ii) Blocks B	$\sqrt{(ln)}$ range $(\bar{x}_{.j.})$	m	—
(iii) Error (1)	$s'_w = \overline{w}'\sqrt{(n)}/c'$	—	ν' (Table II)
(iv) Subtreatment C	$\sqrt{(lm)}$ range $(\bar{x}_{..t})$	n	—
(v) Interaction A × C	$s'_{\mathrm{I}} = \overline{w}'_{\mathrm{I}}\sqrt{(m)}/c'_{\mathrm{I}}$	—	ν'_{I} (Table II)
(vi) Error (2)	$s_w = \overline{w}/c$	—	ν (Table III)

Tests

For treatment A refer (i)/(iii) to tables of q with sample size l and D.F. $= \nu'$.
For treatment C refer (iv)/(vi) to tables of q with sample size n and D.F. $= \nu$.
For error (1) refer (iii)^2/(vi)^2 to tables of F with degrees of freedom ν', ν.

It will be noted that the same analysis is applicable to any three-way cross-classification with one of the three first-order interactions missing in the model, although there may be certain obvious modifications in the tests to be made. Often *a priori* evidence is available allowing us to assume that two of the factors do not interact so that the present model covers an important class of cases.

5. THE THREE-WAY CROSS-CLASSIFICATION

Here the systematic model is given by

$$x_{ijt} = K + A_i + B_j + C_t + I_{ij} + I'_{it} + I''_{jt} + z_{ijt}. \tag{14}$$

Estimation of σ. From (14)

$$y_{ijt} \equiv x_{ijt} - \bar{x}_{.jt} - \bar{x}_{ij.} + \bar{x}_{.j.} = I'_{it} + z_{ijt} - \bar{z}_{.jt} - \bar{z}_{ij.} + \bar{z}_{.j.}. \tag{15}$$

Again we form ranges for fixed i and t giving ln ranges $w_m(i,t)$.

Now $$\text{var}\,(z_{ijt} - \bar{z}_{.jt} - \bar{z}_{ij.} + \bar{z}_{.j.}) = \sigma^2(l-1)(n-1)/(ln).$$

Therefore $$M \equiv \mathscr{E}(\overline{w}) = \sqrt{\left\{\frac{(l-1)(n-1)}{ln}\right\}} d_m \sigma,$$

where d_m has been defined on p. 399 above. Arranging the ranges as in Table 7 ($w_{it} = w_m(i,t)$), we may distinguish three correlation coefficients:

(i) ρ_{w_1} between ranges in the same row,
(ii) ρ_{w_2} between ranges in the same column,
(iii) ρ_{w_3} between ranges neither in the same row or column.

Table 7. *Pattern of ranges of residuals in a three-way cross-classification*

w_{11}	w_{12}			w_{1n}
w_{l1}				w_{ln}

It is easily shown that these are functions respectively of

$$\rho_1 = -1/(n-1), m;\quad \rho_2 = -1/(l-1), m;\quad \rho_3 = 1/\{(n-1)(l-1)\}, m.* \tag{16}$$

Thus $$V \equiv \text{var}\,(\overline{w}) = \frac{(l-1)(n-1)}{l^2n^2}(1 + (n-1)\rho_{w_1} + (l-1)\rho_{w_2} + (n-1)(l-1)\rho_{w_3}) V_m \sigma^2, \tag{17}$$

where V_m has been defined on p. 399 so that V may be obtained by entering Hartley's (1950) table of ρ_w with the three pairs of values of (16).† Hence ν and c can be determined.

* ρ_1 is the correlation between the two 'residuals' y_{ijt_1} and y_{ijt_2} which are in the same (ith) row and correspond to the same j; similarly for ρ_2 and ρ_3.

† This, however, necessitates an extension of the existing table to small positive values of ρ, in view of ρ_3.

It will be realized that the necessity of forming the 'double residuals' y_{ijl} of (15) deprives the range method of much of its speed in the present case. At this stage it may therefore be useful to survey the applicability of this method to the commoner experimental designs. We have seen that a quick procedure is available in all cases when one or more of the interaction parameters of (14) are absent. Moreover, if in an n-way classification ($n > 3$) certain of the higher interactions can be neglected *a priori*, such a simple method will often still suffice. In a replicated 2^n factorial experiment the residual can always be estimated, being the block × treatment interaction, the treatment terms being computed in the usual way so that all squaring of *individual* observations is still avoided. A similar method can be applied to certain cases of confounded factorial designs. The analysis of a Latin square set-up is, however, not possible without the formation of 'double residuals', and these, in addition, have a more complex pattern of correlations than the essentially 'pairwise' correlations which have so far arisen.

6. Simple classification with unequal cell frequencies

We now consider the application of range methods to the analysis of variance of data classified in unequal groups. Both the accuracy of the approximations used and the labour saved (compared with the customary mean square procedure) will in general increase markedly with the total number of observations, which should exceed, say, twenty. In the following example (Kendall, 1946, p. 179) three ways of dealing with the present case are compared: (*a*) the standard technique based on the ratio of two sums of squares; (*b*) a method in which the between-group estimate of variance of (*a*) is divided by a new within-group estimate based on a 'weighted mean range'; and (*c*) the 'unweighted mean range' method in which both the between- and within-group estimates are based on range. The first two are general methods both based on the F-ratio test; the last is valid only in the case of nearly equal cell frequencies, but leads to a particularly short procedure which involves the q-ratio employed by Patnaik for the case of equal cell frequencies.

Table 8. *Lives in hours of four batches of electric lamps*

Batch	Observations	Batch size	Mean	Range
1	1600, 1610, 1650, 1680, 1700, 1720, 1800	7	1680	200
2	1580, 1640, 1640, 1700, 1750	5	1662	170
3	1460, 1550, 1600, 1620, 1640, 1660, 1740, 1820	8	1636	360
4	1510, 1520, 1530, 1570, 1600, 1680	6	1568	170
Total		26		900

Method (*a*). Kendall obtains: between batches mean square = 14,787, residual mean square, $s^2 = 6{,}880$. Hence $s = 82{\cdot}9$ with 22 degrees of freedom.

Method (*b*). The within-batch estimate of error standard deviation is given by the expression

$$s_w = \Sigma w_{ni} \frac{d_{ni}}{V_{ni}} \Big/ \left(\Sigma \frac{d_{ni}^2}{V_{ni}} + \frac{1}{2} \right),$$

with $\nu = \frac{1}{2}\Sigma(d_{ni}^2/V_{ni})$. The quantities d_{ni}/V_{ni} and d_{ni}^2/V_{ni} are listed in Table IV for $n_i = 2-20$.* Thus

$$s_w = (200 \times 3{\cdot}90 + \ldots + 170 \times 3{\cdot}12)/(10{\cdot}54 + \ldots + 8{\cdot}93 + 0{\cdot}5),$$

yielding $s_w = 87{\cdot}4$ with 19·4 degrees of freedom. The numerator of the F-ratio is calculated as in (*a*).

Method (*c*). Here the estimate of residual error is simply $\overline{w}/c = 900/(4 \times 2{\cdot}66) = 84{\cdot}6$ with $\nu = 19{\cdot}7$, c and ν being found by interpolation (at sight) in Table I, entered with $n = \Sigma n_i/4 = 6{\cdot}5$ and $k = 4$. The three analyses may be summarized as shown in Table 9.

Table 9. *Comparison of three methods of analysing a simple classification with unequal cell frequencies*

Method	Value of test ratio	Upper 5 % point
(*a*) Standard	$F_{3,22} = 14{,}787/6{,}880 = 2{\cdot}15$	3·05
(*b*) Weighted mean range	$F_{3,19{\cdot}4} = 14{,}787/(87{\cdot}4)^2 = 1{\cdot}94$	3·12
(*c*) Unweighted mean range	$q_{4,19{\cdot}7} = 112 \times \sqrt{(6{\cdot}5)}/84{\cdot}6 = 3{\cdot}38$	4·0

Thus all three methods show that the differences are insignificant at the 5 % level. More precisely, we find that the three test ratios fall respectively, at the upper 12, 16 and 11 % points. The theoretical bases of methods (*b*) and (*c*) are given in the next two subsections.

(6·1). *The weighted mean-range method*

For the set-up under consideration, namely,

$$x_{ij} = K + A_i + z_{ij} \quad \text{or} \quad x_{ij} = K + u_i + z_{ij} \quad (i = 1 \ldots l;\ j = 1 \ldots n_i),$$

a weighted mean of ranges $\sum_{i=1}^{l} \alpha_i w_{ni}$ (α_i constant) is evidently a suitable and computationally convenient generalization of the mean range $\overline{w}$ so far employed as a measure of variability. To obtain the 'best unbiased estimator' s_w of this form, we solve for the α_i by minimizing $\Sigma\alpha_i^2 V_{ni}$ with respect to the α_i, subject to the condition $\mathscr{E}(\Sigma\alpha_i w_{ni})^2 = \sigma^2$. The usual 'undetermined multiplier' technique leads to

$$s_w = \Sigma\alpha_i w_{ni} = \left(\Sigma w_{ni}\frac{d_{ni}}{V_{ni}}\right) \Big/ \left\{\Sigma\frac{d_{ni}^2}{V_{ni}} + \left(\Sigma\frac{d_{ni}^2}{V_{ni}}\right)^2\right\}^{\frac{1}{2}}$$

$$\doteq \Sigma w_{ni}\frac{d_{ni}}{V_{ni}} \Big/ \left(\Sigma\frac{d_{ni}^2}{V_{ni}} + \frac{1}{2}\right),$$

as $\Sigma d_{ni}^2/V_{ni} \geqslant 1$. Also the χ-approximation gives to sufficient accuracy the degrees of freedom of s_w as $\nu = \frac{1}{2}\Sigma d_{ni}^2/V_{ni}$.†

This solves the problem of estimating σ; if $l = 2$ it therefore enables us to make a t-test for the significance of the difference of two randomly chosen means of, say, n_1 and n_2

* As before d_{ni} and V_{ni} are the expectation and variance of the range in samples of n_i independent unit normal variates.

† Large cells ($n_i \geqslant 12$) may for greater efficiency be randomly subdivided into groups of approximately eight (compare Grubbs & Weaver 1947).

observations, since $(\bar{x}_1-\bar{x}_2)/\{s_w(1/n_1+1/n_2)^{\frac{1}{2}}\}$ follows approximately a t-distribution with ν degrees of freedom. For $l>2$ no simple range statistic seems, in general, available for the estimation of the 'between-groups' term, which must be computed as a weighted sum of squares in the standard way. However, for nearly equal cell frequencies, a special procedure is possible and is developed in the following section.

(6·2). *Range test for a simple classification with approximately equal cell frequencies*

In the case of *equal* cell frequencies ($n_i = n$; $i = 1 \dots l$) Patnaik (1950) has shown that a test for between-group variability is obtained by gauging $q = \sqrt{n}$ range $(\bar{x}_{i.})/(\overline{w}/c)$ against the 100α % point of the 'Studentized' range for l and ν, where both c and ν are given by Table I. It is therefore natural to consider the use of q when the n_i are approximately equal, with $\overline{n}$, the mean cell frequency, replacing n. Some numerical investigation has indicated, as would be expected, that the inaccuracy introduced by this approximation is slight for the denominator of q. We now consider how the numerator is affected and what range of variation of the n_i is permissible for a sufficiently accurate use of this test.

On the null hypothesis ($A_i = 0$ or $u_i = 0$; $i = 1 \dots l$) we have

$$w_l \equiv \text{range}\,(\bar{x}_{i.}) = \text{range}\,(\bar{z}_{i.}) \quad (i = 1 \dots l).$$

To simplify the notation, write $\bar{z}_{i.}$ as z_i. Then the joint probability of the inequalities

$$z_i \leqslant \bar{z}_{i.} \leqslant z_i + dz_i \quad \text{and} \quad z_i \leqslant z_h \leqslant z_i + w \quad (h = 1 \dots l;\ h \neq i)$$

is

$$f_i(z_i) \prod_{h \neq i} \left(\int_{z_i}^{z_i+w} f_h(z_h)\, dz_h \right) dz_i,$$

where

$$f_i(z_i) = (2\pi)^{-\frac{1}{2}} n_i^{\frac{1}{2}} \sigma^{-1} \exp\left[-\frac{1}{2}\frac{n_i}{\sigma^2} z_i^2 \right].$$

This is the probability that the mean of the ith group falls in an element dz_i, that the $l-1$ other group means are all $\geqslant z_i$, and that the range of l means is $\leqslant w$. As the probability density functions f_i are different for each group, to find the probability integral of w we must not only integrate for z_i but also sum for f_i. It follows that the probability integral of w_l is given by

$$P(w \mid n_1, \dots n_l) = \sum_{i=1}^{l} \int_{-\infty}^{\infty} f_i(z_i) \prod_{h \neq i} \left(\int_{z_i}^{z_i+w} f_h(z_h)\, dz_h \right) dz_i. \tag{18}$$

This expression may be regarded formally as a continuous function of the n_i and expressed, by means of a multivariate Taylor expansion, in terms of its value at $n_i = g(n)$ $(i = 1 \dots l)$, where $g(n)$ is some central measure of the n_i. Thus, if $\Delta n_i = n_i - g(n)$, and we write $P(n_i)$ for $P(w \mid n_1, \dots n_l)$, and $P(g)$ for the same expression when all $n_i = g(n)$, we have

$$P(n_i) = P(g) + \sum_{i=1}^{l} \Delta n_i \frac{\partial P(g)}{\partial n_i} + \sum_{i=1}^{l} \sum_{j=1}^{l} \frac{\Delta n_i \Delta n_j}{2!} \frac{\partial^2 P(g)}{\partial n_i \partial n_j} + \dots, \tag{19}$$

where

$$\frac{\partial P(g)}{\partial n_i} = \left(\frac{\partial P(n_i)}{\partial n_i} \right)_{n_1 = g(n),\ \dots,\ n_l = g(n)}, \text{ etc.}$$

Now $P(n_i)$ is symmetric in the n_i. Hence $\partial P(g)/\partial n_i$ has the same value for all i, so that $\sum_{i=1}^{l} \Delta n_i \frac{\partial P(g)}{\partial n_i} = \frac{\partial P(g)}{\partial n_i} \sum_{i=1}^{l} \Delta n_i$, an expression which will vanish if we choose $g(n) = \overline{n}$. Likewise $\partial^2 P(g)/(\partial n_i \partial n_j)$ has only the two different values $\partial^2 P(g)/\partial n_i^2$ and $\partial^2 P(g)/\partial n_i \partial n_h$ $(h \neq i)$.

(19) may now be written

$$P(n_i) = P(\overline{n}) + \frac{1}{2!}\left(\sum_{i=1}^{l}\Delta n_i^2 \frac{\partial^2 P(\overline{n})}{\partial n_i^2} + \sum_{i=1}^{l}\sum_{h=1}^{l}\Delta n_i \Delta n_h \frac{\partial^2 P(\overline{n})}{\partial n_i \partial n_h}\right) + \dots$$

$$= P(\overline{n}) + \frac{1}{2}\left(\frac{\partial^2 P(\overline{n})}{\partial n_i^2}\sum_{i=1}^{l}\Delta n_i^2 + \frac{\partial^2 P(\overline{n})}{\partial n_i \partial n_h}\sum_{i=1}^{l}\sum_{h=1}^{l}\Delta n_i \Delta n_h\right) + \dots$$

$$= P(\overline{n}) + \frac{1}{2}\left(\frac{\partial^2 P(\overline{n})}{\partial n_i^2} - \frac{\partial^2 P(\overline{n})}{\partial n_i \partial n_h}\right)\sum_{i=1}^{l}\Delta n_i^2 + \dots.$$

It is considered that the last term given, which may be denoted by $-C(w \mid \overline{n}, l)\,\Sigma\Delta n_i^2$, accounts essentially for the difference between $P(n_i) = P(w \mid n_1, \dots n_l)$ and our approximation $P(\overline{n}) = P(w \mid \overline{n}, \dots \overline{n})$.* From (18), we can obtain

$$P(\overline{n}) = l\int_{-\infty}^{\infty} z(y)\left(\int_{y}^{y+w\sqrt{n}} z(\eta)\,d\eta\right)^{l-1} dy, \tag{20}$$

where $z(y) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}$. Introduce

$$\phi_l(y, w\sqrt{n}) = z(y)\left(\int_{y}^{y+w\sqrt{n}} z(\eta)\,d\eta\right)^{l-1}$$

and the operator ∇ defined by

$$\nabla(h(y)) = h(y) - h(y + w\sqrt{n})$$

(where $h(y)$ is any function of y). After some heavy algebra it can be shown that

$$8\overline{n}^2 C(w \mid \overline{n}, l) = \int_{-\infty}^{\infty}(1 + 2y^2 - y^4)\,\phi_l(y)\,dy$$

$$+ \int_{-\infty}^{\infty}\{[2y^2 - (l+1)]\,\nabla[yz(y)] - (l-1)\,\nabla[y^3 z(y)]\}\,\phi_{l-1}(y)\,dy$$

$$+ (l-2)\int_{-\infty}^{\infty}(\nabla[yz(y)])^2\,\phi_{l-2}\,dy. \tag{21}$$

In Table 10*a* some typical values of $\overline{n}^2 C(w \mid \overline{n}, l)$ are given, evaluated by approximate numerical quadrature for $w = w_\alpha/\sqrt{n}$, where w_α is the upper $100\alpha\,\%$ point of the range w in samples of l independent unit normal variates. $\overline{n}^2 C(w_\alpha/\sqrt{n} \mid \overline{n}, l)$ is from (21) a function of l and α only and may be termed $C_{l,\alpha}$. The correction to $P(w \mid \overline{n}, \dots \overline{n})$ is always negative, so that the significance level of w is, as a result of the approximation, somewhat lower than $100\alpha\,\%$. An example of the use of Table 10*a* is the construction of upper bounds to $\Sigma\Delta n_i^2/\overline{n}^2$ and to the coefficient of variation of the n_i, supposing that we make the assumption that reductions of the significance level from 95 to 94 % and 99 to 98·5 % are permissible (Table 10*b*). It will be seen that as l increases the coefficient of variation permissible to achieve the prescribed accuracy slightly decreases.

As stated earlier, the use of $\overline{n}$ in place of the actual cell frequencies n_i hardly affects the denominator $\overline{w}/c$ of the q-ratio employed in method (*c*). Table 10*b* will therefore also serve in the present case, which is the 'Studentized' form of the approximate test just examined, and is almost certainly on the safe side. For an illustration, we revert to the example given at the beginning of this section. Here $\overline{n} = 6{\cdot}5$, $l = 4$, $\Sigma\Delta n_i^2 = 5$, so that $\Sigma\Delta n_i^2/\overline{n}^2 = 0{\cdot}12$, a 'permissible' value.

* This has been numerically verified to be a close approximation (for $P(\overline{n})$ equal to 0·95 and 0·99) in the case $l = 2$ when exact values of $P(n_i)$ can be found from tables of the normal integral. It is, of course, realized that $l = 2$ is a rather special case. The third Taylor term of (19) may be shown to vanish if $\sum_{i=1}^{l}\Delta n_i^3 = 0$, and is therefore always zero if $l = 2$.

Table 10*a*. *Values of* $C_{l,\alpha}$ *at upper* 100α *% point*

α \ l	2	4	6	8	10
0·05	0·057	0·053	0·048	0·043	0·038
0·01	0·019	0·019	0·018	0·017	0·015

Note. $P(w|n_1, \dots n_l) = P(w|\bar{n}, \dots \bar{n}) - C_{l,\alpha} \times \Sigma\Delta n_i^2/\bar{n}^2$.

Table 10*b*. *Maximum permissible values of* $\Sigma\Delta n_i^2/\bar{n}^2$ *and (in brackets) of the coefficient of variation of the* n_i, *i.e. of* $\sqrt{(\Sigma\Delta n_i^2/l)}/\bar{n}$

α \ l	2	4	6	8	10
0·05	0·18 (0·30)	0·19 (0·22)	0·21 (0·19)	0·23 (0·17)	0·26 (0·16)
0·01	0·27 (0·37)	0·27 (0·26)	0·28 (0·22)	0·30 (0·19)	0·33 (0·18)

APPENDIX

Use and construction of Tables I–IV

(1) *Table I*. This table giving c and ν appropriate to the mean of k *uncorrelated* ranges in samples of n is an extension of Patnaik's (1950) Table 1, which covered the range $n = 3-10$, $k = 1-5$. It has been obtained by rounding-off. A few corrections for small ν were found necessary. For these, c was calculated from the exact equation

$$c^2 = d_n^2 + V_n/k,$$

and ν (Florin, 1950) from

$$\nu = A^{-1} + \tfrac{1}{4} - \tfrac{3}{16}A + \tfrac{3}{64}A^2,$$

where $A = 2V/M^2$, which gives ν to the required accuracy.

The extension for $k > 5$ is based on equations (9) and (10) from which, for large values of ν,

$$\nu \doteqdot M^2/(2V) \quad (9'), \quad \text{and } c \doteqdot M \quad (10').$$

Writing d_n and V_n for the expectation and variance of the range in samples of n independent unit normal variates, the variance of the mean of k ranges, each based on n observations, is $V = V_n/k$. Thus (9′) is equivalent to

$$\nu = kd_n^2/(2V_n),$$

so that $d_n^2/(2V_n)$ may be regarded as a 'constant' difference (denoted by C.D. in Tables I and II).

Example. Find c and ν for $n = 8$, $k = 12$. We have $c = 2{\cdot}86$, and from the last two entries in the ν-column for $n = 8$ we obtain $\nu = 60{\cdot}6 + 2(6{\cdot}03) = 72{\cdot}7$.

(2) *Table II* is a similar extension of Hartley's (1950) Table 5.

(3) *Table III*. For the split-plot residual error we find from (13) that

$$M^2/V = ln d_m^2/\{V_m[1 + (n-1)\rho_w(n, m)]\}. \qquad (13')$$

The quantities ν and c, which are here functions of l, m and n, may be determined from (9) and (10) as usual. This would necessitate three-way tables which can, however, be avoided by noting that for the double classification of blocks and subtreatments, namely,

$$x_{jt} = K + B_j + C_t + z_{jt} \quad (j = 1 \dots m;\ t = 1 \dots n),$$

we have

$$M^2/V = n d_m^2/\{V_m[1 + (n-1)\rho_w(n, m)]\}, \qquad (22)$$

so that M^2/V in (13′) is l times as large as in (22). It follows from (9′) that a rough value of ν is given by l times the degrees of freedom of the entry in Table II with the same n and $k = m$, and from (10′) that the corresponding c is little changed. In Table III both ν and c of Table II have been slightly adjusted to make the determination of the split-plot error ν and c as accurate as possible for values of l likely to occur in practice.

Example. Take $m = 5$, $n = 3$. From Table III we read off $\nu' = 7{\cdot}0$, $c = 1{\cdot}53$, so that we should take $\nu = 7{\cdot}0l$. Direct calculation gives:

$$\begin{aligned} \text{for } l = 2: &\quad \nu = 14{\cdot}1, \ c = 1{\cdot}54; \\ l = 3: &\quad \nu = 21{\cdot}1, \ c = 1{\cdot}53; \\ l = 5: &\quad \nu = 35{\cdot}0, \ c = 1{\cdot}52. \end{aligned}$$

It may be noted that the corresponding unadjusted values of Table II, namely, $\nu = 7{\cdot}2$, $c = 1{\cdot}57$, would have given much less accurate values for the scale factor and degrees of freedom of the split-plot residual.

(4) *Table IV.* $d_n, d_n/V_n$ and d_n^2/V_n are listed for $n = 2(1)20$. Values of d_n/V_n and d_n^2/V_n have been taken or computed from tables given elsewhere in this issue (Hartley & Pearson, p. 463). d_n has been included to allow the simultaneous computation on a calculating machine of the sums $\Sigma d_n^2/V_n$ and $\Sigma d_n w_n/V_n$, by setting d_n and w_n at the same time and multiplying both by d_n/V_n.

I am greatly indebted to Dr H. O. Hartley for his kind advice and encouragement throughout this investigation and also to Prof. E. S. Pearson for some suggestions as to the improvement of the text.

REFERENCES

DAVIES, O. L. (1947). *Statistical Methods in Research and Production.* London: Oliver and Boyd.
FLORIN, H. (1950). *Communications of the Royal Flemish Academy* (Science Series), **12**, 6.
GRUBBS, F. E. & WEAVER, L. C. (1947). *J. Amer. Statist. Ass.* **42**, 224.
HARTLEY, H. O. (1950). *Biometrika*, **37**, 271.
HARTLEY, H. O. & PEARSON, E. S. (1951). *Biometrika*, **38**, 463.
KENDALL, M. G. (1946). *The Advanced Theory of Statistics*, **2**. London: Griffin.
PATERSON, D. D. (1939). *Statistical Techniques in Agricultural Research.* New York and London: McGraw-Hill.
PATNAIK, P. B. (1950). *Biometrika*, **37**, 78.
PEARSON, E. S. & HARTLEY, H. O. (1943). *Biometrika*, **33**, 89.
SNEDECOR, G. (1946). *Statistical Methods*, 4th ed. Ames, Iowa: Collegiate Press.

Table I. *Scale factor c and equivalent degrees of freedom ν appropriate to the mean of k uncorrelated ranges of n observations*

k \ n	2		3		4		5		6		7		8		9		10	
	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c
1	1·0	1·41	2·0	1·91	2·9	2·24	3·8	2·48	4·7	2·67	5·5	2·83	6·3	2·96	7·0	3·08	7·7	3·18
2	1·9	1·28	3·8	1·81	5·7	2·15	7·5	2·40	9·2	2·60	10·8	2·77	12·3	2·91	13·8	3·02	15·1	3·13
3	2·8	1·23	5·7	1·77	8·4	2·12	11·1	2·38	13·6	2·58	16·0	2·75	18·3	2·89	20·5	3·01	22·6	3·11
4	3·7	1·21	7·5	1·75	11·2	2·11	14·7	2·37	18·1	2·57	21·3	2·74	24·4	2·88	27·3	3·00	30·1	3·10
5	4·6	1·19	9·3	1·74	13·9	2·10	18·4	2·36	22·6	2·56	26·6	2·73	30·4	2·87	34·0	2·99	37·5	3·10
10	9·0	1·16	18·4	1·72	27·6	2·08	36·5	2·34	44·9	2·55	52·9	2·72	60·6	2·86	67·8	2·98	74·8	3·09
d_n		1·13		1·69		2·06		2·33		2·53		2·70		2·85		2·97		3·08
C.D.	0·88		1·82		2·74		3·62		4·47		5·27		6·03		6·76		7·45	

N.B. C.D. = constant difference.

Table II. *Scale factor c and equivalent degrees of freedom ν for analysis of double classification*

k \ n	2		3		4		5		6		7		8		9	
	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c	ν	c
2	1·0	1·00	2·0	1·35	2·9	1·58	3·8	1·75	4·7	1·89	5·5	2·00	6·3	2·10	7·0	2·18
3	1·9	1·05	3·7	1·48	5·6	1·76	7·4	1·96	9·3	2·12	11·3	2·26	13·4	2·37	15·7	2·46
4	2·7	1·07	5·4	1·54	8·2	1·84	11·0	2·06	13·9	2·23	16·9	2·38	20·1	2·50	23·6	2·60
5	3·6	1·08	7·2	1·57	10·9	1·88	14·6	2·12	18·5	2·30	22·4	2·45	26·6	2·57	31·1	2·68
6	4·5	1·09	8·9	1·59	13·6	1·91	18·2	2·15	23·0	2·34	27·9	2·49	33·0	2·62	38·3	2·73
7	5·4	1·09	10·7	1·61	16·3	1·93	21·8	2·18	27·6	2·37	33·3	2·52	39·3	2·65	45·4	2·76
8	6·3	1·10	12·5	1·62	19·0	1·95	25·4	2·20	32·1	2·39	38·7	2·55	45·6	2·68	52·5	2·79
9	7·1	1·10	14·3	1·63	21·7	1·96	29·0	2·21	36·6	2·41	44·0	2·57	51·8	2·70	59·6	2·81
10	8·1	1·10	16·1	1·63	24·4	1·97	32·6	2·22	41·0	2·42	49·3	2·58	57·9	2·71	66·6	2·83
20	16·7	1·11	33·9	1·66	51·5	2·02	68·8	2·28	86·0	2·48	103	2·64	119	2·78	134	2·90
d_n		1·13		1·69		2·06		2·33		2·53		2·70		2·85		2·97
C.D.	0·87		1·80		2·71		3·62		4·50		5·33		6·10		6·79	

Table III. *Scale factor c and equivalent degrees of freedom $\nu(=l\nu')$ for the split-plot design with l main treatments, m blocks and n subtreatments*

m \ n	2		3		4		5		6		7		8		9	
	ν'	c	ν'	c	ν'	c	ν'	c	ν'	c	ν'	c	ν'	c	ν'	c
2	0·9	0·84	1·9	1·23	2·8	1·48	3·7	1·67	4·6	1·81	5·4	1·94	6·1	2·04	6·9	2·13
3	1·8	0·95	3·6	1·40	5·4	1·70	7·2	1·92	9·2	2·08	11·1	2·24	13·2	2·35	15·7	2·45
4	2·6	1·00	5·3	1·48	8·1	1·80	10·9	2·04	13·8	2·22	16·8	2·36	19·9	2·48	23·5	2·59
5	3·5	1·03	7·0	1·53	10·7	1·86	14·5	2·10	18·3	2·28	22·3	2·43	26·5	2·57	30·9	2·67
6	4·4	1·05	8·8	1·56	13·4	1·90	18·1	2·14	22·9	2·33	27·7	2·48	32·8	2·61	38·1	2·72
7	5·2	1·06	10·6	1·59	16·2	1·92	21·7	2·17	27·5	2·36	33·2	2·51	39·1	2·64	45·2	2·76
8	6·1	1·07	12·3	1·60	18·8	1·94	25·2	2·19	32·1	2·38	38·6	2·54	45·3	2·67	52·2	2·79
9	7·0	1·08	14·1	1·61	21·6	1·95	28·8	2·20	36·6	2·40	44·0	2·56	51·6	2·69	59·2	2·81
10	7·9	1·09	15·9	1·62	24·3	1·96	32·5	2·21	41·1	2·41	49·3	2·57	57·7	2·71	66·2	2·82

Table IV. *Weighting factors d_n, d_n/V_n and d_n^2/V_n for analysis of single classification with unequal cell frequencies*

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
d_n	1·13	1·69	2·06	2·33	2·53	2·70	2·85	2·97	3·08	3·17	3·26	3·34	3·41	3·47	3·53	3·59	3·64	3·69	3·73
d_n/V_n	1·55	2·15	2·66	3·12	3·52	3·90	4·24	4·55	4·84	5·12	5·38	5·63	5·87	6·09	6·29	6·50	6·69	6·87	7·03
d_n^2/V_n	1·75	3·63	5·48	7·25	8·93	10·54	12·06	13·51	14·90	16·2	17·5	18·8	20·0	21·1	22·2	23·3	24·3	25·3	26·2

THE FITTING OF POLYNOMIALS TO EQUIDISTANT DATA WITH MISSING VALUES

By H. O. HARTLEY

1. Introduction

It is well known that the fitting of polynomials to observed data $y_t, x_t (t = 1, 2, \ldots, n)$ is considerably simplified if the x_t are equidistantly spaced. Changing scale and origin of x to the standard values

$$x_t \equiv \begin{cases} -\frac{1}{2}(n-1), \ldots, -1, 0, 1, \ldots \frac{1}{2}(n-1) & \text{for } n \text{ odd,} \\ -\frac{1}{2}(n-1), \ldots, -\frac{1}{2}, \frac{1}{2}, \ldots \frac{1}{2}(n-1) & \text{for } n \text{ even,} \end{cases} \tag{1}$$

the orthogonalization of the powers x_t^i leads to the Chebycheff polynomials which have been tabulated by various authors (Anderson & Houseman, 1942; Aitken, 1933; Fisher & Yates, 1948; Van der Reyden, 1943), and considerably simplify the procedure of fitting. Although with observed series the x_t are often basically equidistant, situations frequently arise in which for certain x_t in the series no observed y_t are available, so that these missing y-values spoil the even sequence of the x-series. As examples we may mention here experimental plans providing for observations to be made at regular time intervals in which occasional omissions have occurred, or official statistics in which the interval of tabulation was changed (e.g. doubled or halved) over the time span under consideration. Such situations can, of course, always be dealt with by applying the general technique of polynomial graduation available for non-equidistant x_t. However, it is the purpose of this note to show that in these situations a great deal of computational labour can be saved through treating the data by what may be regarded as a 'missing plot technique' for orthogonal polynomials. It will become apparent that the tables of orthogonal polynomials provided in the literature are eminently suited to the present procedure.

2. Orthogonal fit by estimation of missing values

Let $x_t (t = 1, 2, \ldots, n)$ denote a sequence of equidistant abscissa values which, without loss of generality, are taken in the form (1). Let x_τ denote a subset of n' of the x_t at which observations y_τ are available, and x_θ the complementary set of n'' abscissa values so that $n = n' + n''$. To fit a polynomial $y = P_k(x)$ of degree k to the y_τ we take it in the form

$$P_k(x_t) \equiv \sum_{i=0}^{k} a_i \xi_i'(x_t), \tag{2}$$

where the $\xi_i'(x_t)$ are the familiar orthogonal polynomials for sample size n (e.g. Fisher & Yates, 1947). Under the usual assumptions of normally distributed independent observational errors of equal variance, the a_i are determined as the least squares solution of

$$Q'(a_i) \equiv \sum_{x_\tau} (y_\tau - \sum_i a_i \xi_i'(x_\tau))^2 = \min. \tag{3}$$

The minimization problem (3) can now be modified as follows: simultaneously with the a_i we estimate the ordinate values $y_\theta = u_\theta$ (say) corresponding to the x_θ. Minimizing therefore

$$Q(a_i, u_\theta) \equiv \sum_{x_\tau} (y_\tau - \sum_i a_i \xi_i'(x_\tau))^2 + \sum_{x_\theta} (u_\theta - \sum_i a_i \xi_i'(x_\theta))^2 = \min. \tag{4}$$

as a function of the $k+1+n''$ variables a_i, u_θ, we obtain from $\partial Q/\partial u_\theta = 0$ the conditions $u_\theta - \sum_i a_i \xi_i'(x_\theta) = 0$, so that the equations $\partial Q/\partial a_i = 0$ are identical with the equations $\partial Q'/\partial a_i = 0$ which the a_i must satisfy in the original problem (3). Problem (4) is therefore mathematically equivalent to problem (3) and will be solved in two stages:

(*a*) For any set of *given* u_θ minimize $Q(a_i, u_\theta)$ and so obtain solutions $a_i(u_\theta)$.

(*b*) Substituting these solutions in $Q(a_i, u_\theta)$ obtain $Q(a_i(u_\theta), u_\theta)$ and minimize this as a function of the u_θ. The solutions u_θ of this problem together with the $a_i(u_\theta)$ will be the solutions of (4) and hence of (3).

Now the solutions of (*a*) are clearly the orthogonal polynomial coefficients for the $n = n' + n''$ 'observations' y_τ, u_θ and are given by

$$a_i(u_\theta) = (\sum_{x_\tau} y_\tau \xi_i'(x_\tau) + \sum_{x_\theta} u_\theta \xi_i'(x_\theta))/\sum_{x_l} \xi_i'^2(x_l). \tag{5}$$

The sum of squares $Q(a_i(u_\theta), u_\theta)$ is therefore the error sum of squares of the orthogonal fit (*a*), so that the solutions of problem (*b*) are the ordinates provided by the polynomial at $x = x_\theta$, i.e. by

$$u_\theta = \sum_i a_i(u_\theta)\, \xi_i(x_\theta). \tag{6}$$

By substituting (5) in (6) or (6) in (5) we could obtain linear equations for the u_θ and $a_i(u_\theta)$ respectively. The most expedient way of obtaining the solutions, however, is to solve the system (5) and (6) by iteration, as will be shown in the next section.

3. An example illustrating the computational procedure

In order to illustrate the iteration process we use an example of $n' = 7$ observations covering an x-range of $n = 9$ equidistant x_l values as shown below:

x_l	-4	-3	-2	-1	0	1	2	3	4
y_τ	12	14	.	20	22	.	36	41	50

We show how to fit a cubic to these data, although we realize that in practice longer series would normally be required for such a fit. In columns 1–4 of Table 1 the orthogonal polynomials ξ_0', ξ_1', ξ_2', ξ_3' for sample size $n = 9$, which can be obtained from Fisher & Yates (1948), are shown together with the sums of squares $\sum_{x_l} \xi_i'^2(x_l)$ given at the foot of each column. In column 5 the observed y_τ are entered. The iteration process based on equations (5) and (6) is now as follows:

(i) Estimate trial values for the y in the missing places $x_\theta = -2, 1$ and call these $U_{-2} = 16$ and $U_1 = 29$ (column (6)).

(ii) Using these to complete the series, compute trial values A_i for the orthogonal coefficients from (5) (using $u_\theta = U_\theta$) to obtain the A_i shown in line (11).

(iii) Using these A_i in place of the a_i in (6), compute new values for the u_θ for $x_\theta = -2, 1$ and enter the first corrections $\Delta U_\theta = u_\theta - U_\theta$ in column (7).

(iv) If these are not negligible compute corrections for the A_i from

$$\Delta A_i = \sum_{x_\theta} \Delta U_\theta \xi_i'(x_\theta)/\sum_{x_l} \xi_i'^2(x_l)$$

and enter these in line (12).

(v) If these are not negligible compute second corrections for the U_θ from

$$\delta U_\theta = \sum_i \Delta A_i \xi_i'(x_\theta)$$

and enter these in column (8).

(vi) If these are not negligible compute second corrections for the A_i from

$$\delta A_i = \sum_{x_\theta} \delta U_\theta \xi_i'(x_\theta) / \sum_{x_t} \xi_i'^2(x_t)$$

and enter in line (13), etc.

In the example the $\delta U_\theta, \delta A_i$ need not have been computed.

Table 1. *Fitting of cubic to equidistant series with two missing values*

Line	Column							
	(1) ξ_0'	(2) ξ_1'	(3) ξ_2'	(4) ξ_3'	(5) y_τ	(6) U_θ	(7) ΔU_θ	(8) δU_θ
(1)	1	−4	28	−14	12	—	—	—
(2)	1	−3	7	7	14	—	—	—
(3)	1	−2	− 8	13	—	16	0·05	0·01
(4)	1	−1	−17	9	20	—	—	—
(5)	1	0	−20	0	22	—	—	—
(6)	1	1	−17	− 9	—	29	−0·30	−0·10
(7)	1	2	− 8	−13	36	—	—	—
(8)	1	3	7	− 7	41	—	—	—
(9)	1	4	28	14	50	—	—	—
(10) $\sum_{x_t} \xi_i'^2$	9	60	2772	990	—	—	—	—
(11) A_i	26·67	4·700	0·156	0·002	—	—	—	—
(12) ΔA_i	− 0·03	−0·007	0·002	0·003	—	—	—	—
(13) δA_i	− 0·01	−0·002	0·001	0·001	—	—	—	—
(14) a_i	26·63	4·691	0·159	0·006	—	—	—	—

The final polynomial coefficients a_i are obtained from $a_i = A_i + \Delta A_i + \delta A_i$ and are shown in line (14).

4. The test of significance

Since the a_i determined in § 2 are the correct solutions of the least square problem (3), the $k+1$ degrees of freedom sum of squares corresponding to the fitted polynomial terms of degree 0 to k is given by

$$\sum_{i=0}^{k} a_i \sum_{x_\tau} y_\tau \xi_i'(x_\tau).$$

As a matter of computational convenience, therefore, one would record the sums $\sum_{x_\tau} y_\tau \xi_i'(x_\tau)$ separately when computing the A_i in stage (ii) of § 3. In the example these sums are as follows:

$$\Sigma y_\tau = 195, \quad \Sigma y_\tau \xi_1' = 285, \quad \Sigma y_\tau \xi_2' = 1053, \quad \Sigma y_\tau \xi_3' = 55.$$

The resulting analysis of variance table is set out below:

	Formulae		Example	
	Sum of squares	D.F.	Sum of squares	D.F.
Total (including correction for mean)	Σy_τ^2	n'	6701	7
Polynomial terms	$\sum_i a_i \sum_{x_\tau} y_\tau \xi_i'(x_\tau)$	$k+1$	6698	4
Error	$\sum_{x_\tau} (y_\tau - \sum_i a_i \xi_i'(x_\tau))^2$	$n'-k-1$	3	3

As an approximate analysis of variance one could carry out the familiar computation of constant, linear, quadratic, cubic, ... terms on the completed series of n terms y_τ, u_θ, but reduce the degrees of freedom for 'Total' and 'Error' to n' and $n'-k-1$ respectively. Both the exact and particularly the approximate analysis of variance can conveniently be carried out sequentially by fitting and testing in turn the various polynomial terms in the order of their hierarchy.

REFERENCES

AITKEN, A. C. (1933). *Proc. Roy Soc. Edinb.* **53**, 54.

ANDERSON, R. L. & HOUSEMAN, E. E. (1942). *Res. Bull.* no. 297, Agric. Exp. Station, Iowa State College, Ames.

FISHER, R. A. & YATES, F. (1948). *Statistical Tables*, 3rd ed. Oliver and Boyd.

VAN DER REYDEN, D. (1943). *Onderstepoort J. Vet. Sci.* **18**, 355.

TRUNCATED LOGNORMAL DISTRIBUTIONS

I. SOLUTION BY MOMENTS

By H. R. THOMPSON

Applied Mathematics Laboratory, Department of Scientific and Industrial Research, Wellington, New Zealand

1. Introduction

An important contribution to the study of skew-frequency curve problems was made when Spiller (1948) showed that many discrete distributions of J-shaped and related forms could be normalized by a logarithmic transformation of the enumeration variable. Using the device of plotting the cumulative distribution of the observations as probits against a logarithmic scale of counts, he found that the distributions became straight lines even though sometimes up to 70 % of the data was congregated at the first point. Both Spiller and the present writer have since found numerous distributions amenable to the same treatment, in all of which the variable is either a count or an occurrence. Examples of the former are counts of organisms (fungal spores on culture plates) or insects (aphides on leaves); of the latter, the frequency of accidents to bus drivers quoted by Yule (1944, table 3·2) or the number of words in sentences from different authors (Williams, 1940). Since in these discrete distributions we have a discontinuous variate which can take only integral values from 0 to infinity, it is obvious that the frequency of r counts per sampling unit must be represented by the equation

$$y_r = \frac{1}{\sigma\sqrt{(2\pi)}} \int_{\log r}^{\log (r+1)} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx, \quad \sum_{r=0}^{\infty} y_r = 1, \tag{1·1}$$

where $\log (r+1)$ is the normalizing function (all logarithms to base 10) and μ, σ are parameters of the parent normal distribution. In a lognormal distribution with continuous variate, the curve (after transformation) is recognizably normal throughout. In the discrete case, however, the number appearing as zeros represents the integral over the range from $-\infty$ to 0, giving a markedly skew appearance to the curve. This, combined with the fact that negative values cannot explicitly occur, led to the descriptive name 'truncated lognormal distribution'.

The inadequacy of some attempted methods of representing this type of distribution by other theoretical models is well shown in Upholt & Craig (1940), whose data on black scale on citrus twigs is fitted very well by a truncated lognormal (see Spiller). The expected Poisson and Neyman Type C contagious distributions calculated by Upholt & Craig give absurdly high values of χ^2, whereas a truncated lognormal as in (1·1) with $\mu = 0{\cdot}439$, $\sigma = 0{\cdot}609$, gave $\chi^2 = 11{\cdot}3$ for 7 degrees of freedom ($P \doteqdot 0{\cdot}13$), indicating a very reasonable agreement. A comparison of observed and expected values is given in the table on p. 415.

Bliss (1937) observed instances of continuous truncated lognormal distributions in time-mortality data, truncation being either artificial (due to stopping the experiments before mortality was complete) or inherent (due to non-reaction of individuals). The data consist of n individual observations with ordinary normal probability $\frac{1}{\sigma\sqrt{(2\pi)}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

($x \geqslant a$, the point of truncation) and $N-n$ with total probability $\frac{1}{\sigma\sqrt{(2\pi)}}\int_{-\infty}^{a}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$, which lie somewhere in the truncated portion $x < a$. Bliss solved the distribution by maximum likelihood, computing corrections to a preliminary graphic estimate of the parameters by an iterative process. The calculations entailed in this method are fairly involved, and could only be reduced by the use of an extensive bivariate table giving the point of truncation in terms of the fraction in the truncated portion and some function of the first two moments. For such a table see Hald (1949), the distribution being termed by him 'censored'.

No. of scales per unit sample	0	1	2	3	4	5	6	7	8	Over 8
Observed frequency of samples	199	124	106	65	42	46	36	30	14	159
Expected lognormal distribution	193	144	94	66	49	38	30	25	20	163

Gjeddebaek (1949) considered the maximum likelihood solution of the discrete normal distribution described (using our notation) by the equation

$$q_i = \frac{1}{\sigma\sqrt{(2\pi)}}\int_{x_i}^{x_{i+1}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx,$$

of which (1·1) may be regarded as a special case. The beauty of this method lies in the simplicity of the equations of maximum likelihood. If n_i observations of the total sample of n fall in the ith class interval (x_i, x_{i+1}) their probability is q_i from above, and from the likelihood function of the observations the equations for estimating μ, σ respectively reduce to $\Sigma n_i\frac{z_i - z_{i+1}}{q_i} = 0$, $\Sigma n_i\frac{t_i z_i - t_{i+1}z_{i+1}}{q_i} = 0$, where $t_i = \frac{x_i-\mu}{\sigma}$, $z_i = \frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}t_i^2}$. By means of special tables giving $(z_i - z_{i+1})/q_i$, $(t_i z_i - t_{i+1}z_{i+1})/q_i$ for values of $t_i = x$, $t_{i+1} = x+y$, Gjeddebaek reduces these to the computational form $\Sigma n_i z_{1i} = 0$, $\Sigma n_i z_{2i} = 0$, in which case the variance matrix is also very simply expressed:

$$\text{var}\,(\mu, \sigma) = \frac{\sigma^2}{n}\begin{pmatrix}\Sigma q_i z_{1i}^2 & \Sigma q_i z_{1i}z_{2i}\\ \Sigma q_i z_{1i}z_{2i} & \Sigma q_i z_{2i}^2\end{pmatrix}^{-1}.$$

However, the actual mechanics of the solution are even more laborious than Bliss's method, the process again being iterative; when there are more than say four or five class intervals the method becomes almost impracticable.

The essential requirement of a practical method is that it should be reasonably easy to apply, avoiding the laborious calculations inherent in iterative methods. The technique described in the present paper estimates by the method of moments the parameters μ, σ of the discrete frequency distribution (1·1) by an approximation to the equivalent continuous truncated lognormal distribution. Corrections to these estimates have to be applied owing to the unequal size of the logarithmic intervals from which the moments are calculated. An examination of the efficiency of the moment solution compared with that of maximum likelihood will be made in a later paper. In results so far obtained, the comparison is not unfavourable.

2. Solution

Consider a continuous random variable x, distributed normally with mean μ, standard deviation σ, but truncated at $x = a$ so that the frequency function is of the form

$$\left.\begin{aligned} y &= 0 \quad (x<a) \\ &= \frac{1}{\sigma\sqrt{(2\pi)}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (x>a), \end{aligned}\right\} \tag{2·1}$$

with a discrete mass point equal to

$$\frac{1}{\sigma\sqrt{(2\pi)}}\int_{-\infty}^{a}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx \quad \text{at} \quad x=a.$$

Hence $\int_{-\infty}^{\infty} y\,dx = 1$. The assumption that the truncated frequency is accumulated at the point of truncation follows from the consideration that in the discrete frequency distribution which forms the basis of the paper where the truncation is at $x = 0$, negative values of the enumeration variable cannot occur.

This distribution may be solved for μ, σ by equating the sample moments μ_1', μ_2' to those of the theoretical distribution. Taking moments of (2·1) about the point of truncation a, we obtain easily the following expressions:

$$\left.\begin{aligned} \mu_1'(a) &= (z-tq)\,\sigma, \\ \mu_2'(a) &= (q-tz+t^2q)\,\sigma^2, \end{aligned}\right\} \tag{2·2}$$

where
$$t = \frac{a-\mu}{\sigma}, \quad z = \frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}t^2}, \quad q = \int_t^{\infty} z\,dt. \tag{2·3}$$

Write
$$\phi_2 = \frac{(z-tq)^2}{q-tz+t^2q} = \frac{\{\mu_1'(a)\}^2}{\mu_2'(a)}, \tag{2·4}$$

$$\phi_1 = z-tq. \tag{2·5}$$

Then ϕ_2 is independent of σ and is a function of only one unknown, t. Table 1 gives ϕ_2 for t, proceeding by intervals of 0·1 from $+2·0$ to $-4·0$; from this table t may be obtained by inverse interpolation if ϕ_2 is given. The values of μ, σ may then be determined with the help of ϕ_1, also tabulated against t in Table 1. Thus for a sample of n observations of x, from the population (2·1), the parameters μ, σ may be estimated as follows with the minimum of computation:

(i) From the sample moments

$$\mu_1'(a) = \Sigma(x-a)/n, \quad \mu_2'(a) = \Sigma(x-a)^2/n,$$

estimate ϕ_2 from (2·4).

(ii) Hence obtain t from Table 1.

(iii) Using ϕ_1 from Table 1, calculate

$$\sigma = \mu_1'(a)/\phi_1, \quad \mu = a-\sigma t.$$

For $t < -4·0$, the population may be regarded as non-truncated, in which case the parameters are estimated in the usual way from $\mu = \Sigma(x)/n = \bar{x}$, $\sigma^2 = \Sigma(x-\bar{x})^2/n$.

Table 1. *Functions ϕ_1, ϕ_2 of the normal distribution*

t	ϕ_2	Δ_1	Δ_2	ϕ_1	Δ_1	Δ_2	t	ϕ_2	Δ_1	Δ_2	ϕ_1	Δ_1	Δ_2
2·0	0·0125			0·0085			−1·0	0·6098			1·0833		
		33			26				255			853	
1·9	0·0158		7	0·0111		7	−1·1	0·6353		−12	1·1686		22
		41			32				243			875	
1·8	0·0199		8	0·0143		8	−1·2	0·6595		−13	1·2561		19
		49			40				230			894	
1·7	0·0248		10	0·0183		9	−1·3	0·6826		−13	1·3455		17
		59			50				217			911	
1·6	0·0307		11	0·0232		11	−1·4	0·7043		−13	1·4367		15
		69			61				204			926	
1·5	0·0376		12	0·0293		13	−1·5	0·7247		−13	1·5293		13
		81			74				191			939	
1·4	0·0457		13	0·0367		15	−1·6	0·7438		−13	1·6232		11
		94			89				178			950	
1·3	0·0551		14	0·0455		17	−1·7	0·7616		−13	1·7183		9
		108			106				166			960	
1·2	0·0659		15	0·0561		19	−1·8	0·7782		−12	1·8143		8
		123			125				153			968	
1·1	0·0782		16	0·0686		22	−1·9	0·7935		−12	1·9111		7
		139			147				142			974	
1·0	0·0921		16	0·0833		25	−2·0	0·8077		−11	2·0085		5
		155			171				131			980	
0·9	0·1077		17	0·1004		27	−2·1	0·8208		−10	2·1065		4
		172			198				121			984	
0·8	0·1249		17	0·1202		29	−2·2	0·8329		−10	2·2049		4
		189			227				111			988	
0·7	0·1438		17	0·1429		31	−2·3	0·8440		− 9	2·3037		3
		206			258				102			991	
0·6	0·1644		16	0·1687		33	−2·4	0·8542		− 8	2·4027		2
		222			291				94			993	
0·5	0·1866		15	0·1978		35	−2·5	0·8636		− 8	2·5020		2
		238			326				86			995	
0·4	0·2104		14	0·2304		37	−2·6	0·8722		− 7	2·6015		1
		252			363				79			996	
0·3	0·2356		13	0·2668		38	−2·7	0·8801		− 6	2·7011		1
		265			401				73			997	
0·2	0·2621		11	0·3069		39	−2·8	0·8874		− 6	2·8008		1
		276			440				67			998	
0·1	0·2897		9	0·3509		40	−2·9	0·8941		− 5	2·9005		1
		286			480				62			999	
0·0	0·3183		7	0·3989		40	−3·0	0·9002		− 5	3·0004		
		293			520				57			999	
−0·1	0·3476		5	0·4509		40	−3·1	0·9059		− 4	3·1003		
		299			560				52			999	
−0·2	0·3775		3	0·5069		39	−3·2	0·9111		− 4	3·2002		
		302			599				48			999	
−0·3	0·4077		1	0·5668		38	−3·3	0·9160		− 4	3·3001		
		302			637				45			1000	
−0·4	0·4379		− 1	0·6304		37	−3·4	0·9204		− 3	3·4001		
		301			674				41				
−0·5	0·4680		− 3	0·6978		35	−3·5	0·9246		− 3	3·5001		
		298			709				38				
−0·6	0·4978		− 5	0·7687		33	−3·6	0·9284		− 3	3·6000		
		292			742				36				
−0·7	0·5270		− 7	0·8429		31	−3·7	0·9319		− 3	3·7000		
		285			773				33				
−0·8	0·5555		− 9	0·9202		29	−3·8	0·9352		− 2	3·8000		
		276			802				31				
−0·9	0·5831		−10	1·0004		27	−3·9	0·9383		− 2	3·9000		
		266			829				29				
−1·0	0·6098		−11	1·0833		25	−4·0	0·9412		− 2	4·0000		

Table 3. *Values of* $x = \frac{1}{2}[\log r + \log(r+1)]$ *and* x^2

For $r > 100$, the error involved in making the approximation $x = \log(r+\frac{1}{2})$ is negligible.

r	x	x^2	r	x	x^2	r	x	x^2	r	x	x^2	r	x	x^2
1	0·151	0·023	21	1·332	1·775	41	1·618	2·618	61	1·789	3·200	81	1·911	3·652
2	0·389	0·151	22	1·352	1·828	42	1·628	2·652	62	1·796	3·225	82	1·916	3·673
3	0·540	0·291	23	1·371	1·880	43	1·638	2·685	63	1·803	3·250	83	1·922	3·693
4	0·651	0·423	24	1·389	1·930	44	1·648	2·717	64	1·810	3·274	84	1·927	3·713
5	0·739	0·545	25	1·406	1·978	45	1·658	2·749	65	1·816	3·299	85	1·932	3·732
6	0·812	0·659	26	1·423	2·025	46	1·667	2·780	66	1·823	3·323	86	1·937	3·752
7	0·874	0·764	27	1·439	2·071	47	1·677	2·811	67	1·829	3·346	87	1·942	3·771
8	0·929	0·862	28	1·455	2·116	48	1·686	2·842	68	1·836	3·370	88	1·947	3·791
9	0·977	0·955	29	1·470	2·160	49	1·695	2·872	69	1·842	3·393	89	1·952	3·810
10	1·021	1·042	30	1·484	2·203	50	1·703	2·901	70	1·848	3·416	90	1·957	3·828
11	1·060	1·124	31	1·498	2·245	51	1·712	2·930	71	1·854	3·438	91	1·961	3·847
12	1·097	1·202	32	1·512	2·286	52	1·720	2·959	72	1·860	3·461	92	1·966	3·866
13	1·130	1·277	33	1·525	2·326	53	1·728	2·987	73	1·866	3·483	93	1·971	3·884
14	1·161	1·348	34	1·538	2·365	54	1·736	3·015	74	1·872	3·505	94	1·975	3·902
15	1·190	1·416	35	1·550	2·403	55	1·744	3·042	75	1·878	3·527	95	1·980	3·920
16	1·217	1·482	36	1·562	2·441	56	1·752	3·070	76	1·884	3·548	96	1·985	3·938
17	1·243	1·545	37	1·574	2·477	57	1·760	3·096	77	1·889	3·569	97	1·989	3·956
18	1·267	1·605	38	1·585	2·514	58	1·767	3·123	78	1·895	3·590	98	1·993	3·974
19	1·290	1·664	39	1·597	2·549	59	1·775	3·149	79	1·900	3·611	99	1·998	3·991
20	1·312	1·720	40	1·607	2·584	60	1·782	3·175	80	1·906	3·632	100	2·002	4·009

Table 2A. *Corrections to σ′*

μ' \ σ'	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
−0·8								0·0215 13 6	0·0166 4	0·0130
−0·7							0·0252 17 6	0·0192 11 5	0·0149 4	0·0117
−0·6						0·0298 20 9	0·0224 16 7	0·0170 10 4	0·0132 4	0·0104
−0·5					0·0358 31 12	0·0262 18 9	0·0197 13 6	0·0150 10 5	0·0116 4	0·0092
−0·4				0·0414 19 7	0·0311 26 14	0·0227 15 8	0·0171 12 7	0·0130 9 5	0·0101 3	0·0081
−0·3			0·0452 −4 −15 −8	0·0360 26 13	0·0264 22 13	0·0194 14 9	0·0146 10 6	0·0112 7 4	0·0088 3	0·0071
−0·2		0·0255	0·0403 17 −9 8	0·0303 26 14	0·0220 18 12	0·0163 11 8	0·0124 8 6	0·0096 5 4	0·0076 3	0·0061
−0·1		0·0367	0·0339 22 −3 16	0·0247 24 3 18	0·0178 14 4 11	0·0133 9 3 8	0·0102 6 5	0·0080 4 4	0·0064 3	0·0052
0·0		0·0350 8 −49 13	0·0266 25 3 27	0·0190 17 6 19	0·0139 11 3 10	0·0105 6 3 7	0·0082 5 5	0·0065 3 3	0·0053	0·0044
0·1		0·0256 22 1 65	0·0185 17 9 30	0·0136 8 7 16	0·0104 8 5 10	0·0080 4 4 7	0·0064 3 4	0·0052 2	0·0043	0·0036
0·2		0·0113 −13 65 73	0·0107 3 18 28	0·0088 2 8 14	0·0072 5 5 9	0·0058 3 4	0·0049 3	0·0041 3	0·0034	0·0029
0·3		−0·0029 −58 76 26	0·0038 −11 21 18	0·0047 −3 8 11	0·0045 4 7	0·0040 4	0·0035 2	0·0030	0·0026	0·0023
0·4		−0·0106 −66 29 −29	−0·0013 −18 17 5	0·0014 −7 10 7	0·0023 5 5	0·0025 3	0·0024 3	0·0021	0·0019	0·0018
0·5	−0·0097 74 −17 −23	−0·0107 −32 2 −41	−0·0043 −16 15 −2	−0·0011 −9 7 2	0·0005 −5 4 2	0·0012 2	0·0014 3	0·0014	0·0014	0·0014
0·6	−0·0046 56 −25 −4	−0·0079 7 −9 −28	−0·0056 −12 8 −10	−0·0026 −9 6 −1	−0·0008 −5 4 1	0·0001	0·0005	0·0008	0·0009	0·0010
0·7	−0·0012 32 −8 12	−0·0049 25 −8 −11	−0·0054 −3 1 −12	−0·0034 −7 5 −3	−0·0017 −5	−0·0007	−0·0002	0·0002	0·0004	0·0006
0·8	−0·0003 9 0 12	−0·0028 24 −6 −2	−0·0044 6 −2 −8	−0·0036 −5 1 −6	−0·0022 −3	−0·0013	−0·0007	−0·0002	0·0000	0·0003
0·9	−0·0002 −5 6	−0·0015 18 −4 2	−0·0033 8 −5	−0·0033 1 −4	−0·0025	−0·0016	−0·0010	−0·0006	−0·0003	0·0000
1·0	−0·0001 −9 3	−0·0008 11 4	−0·0024 9 −2	−0·0029 3 −3	−0·0025	−0·0018	−0·0012	−0·0008	−0·0005	−0·0003
1·1	−0·0001 −8	−0·0005 5 6	−0·0017 8	−0·0024 4	−0·0023	−0·0018	−0·0013	−0·0010	−0·0007	−0·0005
1·2	−0·0002 −3	−0·0005 −2	−0·0011 7	−0·0019 4	−0·0020	−0·0017	−0·0013	−0·0011	−0·0008	−0·0007

The figures in smaller type are second differences; for key, see (2·6)–(2·8), p. 420.

Table 2B. *Corrections to μ'*

μ' \ σ'	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
−0·8								−0·0395 −37 −5 −19	−0·0286 −3 −11	−0·0214 −3
−0·7							−0·0472 −48 −8 −25	−0·0334 −29 −6 −17	−0·0244 −4 −11	−0·0183 −3
−0·6						−0·0565 −61 −14 −34	−0·0391 −40 −10 −23	−0·0278 −23 −7 −15	−0·0205 −4 −10	−0·0155 −3
−0·5					−0·0690 −92 −15 −49	−0·0458 −59 −13 −30	−0·0318 −32 −9 −20	−0·0228 −18 −6 −12	−0·0170 −4 −9	−0·0130 −2
−0·4				−0·0834 −103 −16 −59	−0·0548 −73 −20 −47	−0·0365 −46 −12 −26	−0·0255 −24 −8 −17	−0·0185 −15 −5 −10	−0·0139 −4 −7	−0·0108
−0·3			−0·0984 −109 −11 −66	−0·0648 −91 −23 −61	−0·0421 −64 −17 −39	−0·0285 −31 −10 −22	−0·0201 −17 −7 −14	−0·0148 −12 −4 −9	−0·0112 −5	−0·0088
−0·2		−0·0856	−0·0750 −104 −28 −78	−0·0480 −69 −26 −55	−0·0314 −35 −16 −32	−0·0217 −23 −10 −19	−0·0155 −12 −7 −11	−0·0116 −8 −4 −7	−0·0089 −5	−0·0070
−0·1		−0·0787 −68 23 −59	−0·0527 −81 −35 −76	−0·0335 −46 −26 −45	−0·0224 −22 −16 −26	−0·0159 −15 −9 −16	−0·0116 −8 −5 −8	−0·0088 −6 −5 −7	−0·0068 −5	−0·0054
0·0		−0·0533 −85 −51 −117	−0·0332 −50 −43 −67	−0·0216 −27 −23 −35	−0·0150 −12 −13 −19	−0·0111 −7 −6 −12	−0·0084 −7 −5 −8	−0·0064 −4 −3 −5	−0·0051 −4	−0·0042
0·1		−0·0256 −35 −101 −109	−0·0172 −18 −43 −47	−0·0123 −11 −23 −25	−0·0092 −5 −13 −14	−0·0072 −3 −8 −9	−0·0057 −4 −5 −6	−0·0045 −3 −4 −4	−0·0037 −3	−0·0032
0·2		−0·0030 27 −101 −51	−0·0055 4 −38 −27	−0·0053 −19 −15	−0·0047 −11 −9	−0·0041 −7 −6	−0·0035 −5 −5	−0·0029 −3 −3	−0·0025	−0·0023
0·3		0·0095 51 −49 12	0·0019 16 −26 −8	−0·0006 6 −14 −7	−0·0015 3 −8 −5	−0·0018 −5 −4	−0·0018 −3 −3	−0·0017	−0·0016	−0·0016
0·4		0·0119 31 −5 33	0·0055 17 −16 4	0·0022 8 −9	0·0006 4 −6	−0·0002 −4	−0·0006	−0·0008	−0·0009	−0·0010
0·5	0·0078 −47 14 2	0·0094 2 6 26	0·0065 12 −5 11	0·0036 7 −7 2	0·0019 4 −4	0·0009 −3	0·0003	−0·0001	−0·0003	−0·0004
0·6	0·0046 −23 7 −6	0·0064 −13 9 13	0·0059 3 −2 9	0·0041 5 −3 5	0·0026 3 −4	0·0016 −3	0·0009	0·0004	0·0002	0·0000
0·7	0·0028 −4 5 −4	0·0040 −17 6 2	0·0048 −1 2 8	0·0039 4	0·0029 2	0·0020	0·0013	0·0008	0·0005	0·0003
0·8	0·0017 2 2 −3	0·0025 −11 3 −2	0·0035 −5 3 5	0·0034 3	0·0028 3	0·0021	0·0015	0·0011	0·0008	0·0006
0·9	0·0011 3 −2	0·0016 −4 −2	0·0024 −7 3 2	0·0028 −1 3	0·0025 3	0·0021	0·0016	0·0012	0·0010	0·0008
1·0	0·0007 3	0·0010	0·0016 −6	0·0022 −1	0·0022	0·0021	0·0017	0·0013	0·0011	0·0010
1·1	0·0005 4	0·0006	0·0011 −3	0·0016	0·0018	0·0019	0·0016	0·0013	0·0012	0·0011
1·2	0·0003	0·0004	0·0007	0·0011	0·0014	0·0015	0·0014	0·0013	0·0012	0·0012

The figures in smaller type are second differences; for key, see (2·6)–(2·8), p. 420.

The extension to the solution of the discrete distribution (1·1) is made by calculating the parameters as if the distribution were in fact continuous, and applying corrections to the estimates for the discontinuities caused by grouping. In the case of a continuous variable grouped into intervals of equal size and with contact of high order with the x-axis at both terminals, the corrections to the raw or grouped moments are functions only of the size of the grouping interval. For a distribution such as (1·1), however, regarded as being derived from the continuous case, there is a lack of high contact at the lower terminal 0 and, in addition, the grouping intervals, being logarithmic, are of unequal size. The unequal class intervals would in themselves lead to very complicated correction formulae, but the lack of high contact means that the Euler-Maclaurin formula on which Sheppard's corrections depend is invalidated. This follows since for the normal function $e^{-\frac{1}{2}x^2}$ the necessary condition that derivatives of all orders vanish at the terminal does not hold; they do not even converge when the terminal frequency is finite as in a truncated distribution. Any modification of the Sheppard corrections such as Hartley's (1950) based on the Gregory integration formula is also inapplicable generally for a grouped truncated normal.

If then for a theoretical distribution with given μ, σ we calculate moments using the mid-points of successive logarithmic intervals and then estimate μ, σ as if the distribution were continuous, we shall obtain biased estimates μ', σ' say, in error by amounts depending on the values of μ, σ. Since the magnitudes of these errors cannot be determined as explicit functions of μ and σ or the first two moments, they have been obtained from a series of known discrete lognormal distributions with $n = 10{,}000$ (Fisher & Yates, *Statistical Tables*, table IX), varying μ and σ so as to cover the range of μ', σ' given in Table 2. The table has been extended to include low values of σ' and high negative values of μ' (the latter corresponding to 80 % or more zeros), although they will seldom if ever be needed, and the corrections are tabulated for interest only. Over the rest of the table, where the numerous experimental distributions studied by the author have fallen, second difference interpolation will give an accuracy commensurate with that of the table itself. For this purpose, second differences for each entry have been tabulated where necessary (i.e. omitted if they have no influence) as shown below:

$$\left.\begin{array}{lll} f(\mu',\sigma') & \Delta^2_{\sigma'\sigma'} & f(\mu',\sigma'+0{\cdot}1), \\ \quad \Delta^2_{\mu'\mu'} & \Delta^2_{\mu'\sigma'} & \\ f(\mu'+0{\cdot}1,\sigma') & & f(\mu'+0{\cdot}1,\sigma'+0{\cdot}1), \text{ etc.,} \end{array}\right\} \tag{2·6}$$

where
$$\Delta^2_{\mu'\mu'} = f(\mu',\sigma') + f(\mu'+0{\cdot}2,\sigma') - 2f(\mu'+0{\cdot}1,\sigma'), \tag{2·7}$$

and similarly for $\Delta^2_{\sigma'\sigma'}$, while

$$\Delta^2_{\mu'\sigma'} = f(\mu',\sigma') + f(\mu'+0{\cdot}1,\sigma'+0{\cdot}1) - f(\mu'+0{\cdot}1,\sigma') - f(\mu',\sigma'+0{\cdot}1). \tag{2·8}$$

A suitable formula for interpolation (where $0 \leqslant m, n < 1$) is then

$$f(\mu'+0{\cdot}1m,\sigma'+0{\cdot}1n) = f(\mu',\sigma') + m\Delta_{\mu'} + n\Delta_{\sigma'} + mn\Delta^2_{\mu'\sigma'} - \tfrac{1}{2}m(1-m)\Delta^2_{\mu'\mu'} - \tfrac{1}{2}n(1-n)\Delta^2_{\sigma'\sigma'}, \tag{2·9}$$

from which terms may be omitted if they do not affect the final result. Often sufficient accuracy will be obtained with simple linear interpolation.

3. Examples

The method of solution of the discrete lognormal distribution is illustrated by two examples from biological sources. The first is a record of counts of penicillia spores observed on Petri

Example 1. Spores of penicillia on 506 Petri dishes.

No. of spores per dish (r)	Observed frequency	x	x^2
0	277	—	—
1	117	0·151	0·023
2	42	0·389	0·151
3	22	0·540	0·291
4	15	0·651	0·423
5	12	0·739	0·545
6	8	0·812	0·659
7	2	0·874	0·764
8	3	0·929	0·862
9	2	0·977	0·955
10	2	1·021	1·042
11	1	1·060	1·124
18	1	1·267	1·605
22	1	1·352	1·828
28	1	1·455	2·116
Total	506		

Example 2. Aphides on forty leaves of variety Osborne Park.

No. of aphides per leaf (r)	Observed frequency	x	x^2
0	9	—	—
1	11	0·151	0·023
2	4	0·389	0·151
3	5	0·540	0·291
5	2	0·739	0·545
7	1	0·874	0·764
8	2	0·929	0·862
10	2	1·021	1·042
11	1	1·060	1·124
27	1	1·439	2·071
43	1	1·638	2·685
222	1	2·347	5·510
Total	40		

Statistic	Example 1	Example 2
μ'_1	0·16735	0·4663
μ'_2	0·09560	0·4841
ϕ_2	0·2930	0·4492
t	0·0885	−0·4375
ϕ_1	0·3562	0·6553
σ'	0·4698	0·7116
μ'	−0·0416	0·3113
Corrections to σ'	0·0171	0·0033
Corrections to μ'	−0·0199	−0·0017
σ	0·4869	0·7149
μ	−0·0615	0·3096

	Example 1								Example 2					
Class interval	0	1	2	3	4	5	6–7	8 and over	0	1	2	3–4	5–9	Over 9
Observed frequency	277	117	42	22	15	12	10	11	9	11	4	5	5	6
Expected frequency	278·4	112·1	47·5	24·2	13·8	8·6	9·4	12·0	13·3	6·5	3·9	4·6	5·0	6·7

$\chi^2_{[5]} = 2{\cdot}62$, $P = 0{\cdot}76$ $\chi^2_{[5]} = 4{\cdot}61$, $P = 0{\cdot}20$

dishes exposed to the atmosphere for 5 min. at six localities in Auckland over a period of 12 months. The second represents counts of the common carrot aphid (*Cavariella aegopodii*), part of a varietal infestation trial on carrots. Samples of forty leaves were taken, the results

for one variety being reproduced here. The successive steps in computing the mean and standard deviation of the parent distributions are set out in a convenient form in the tables on p. 421. For the calculation of moments, the distributions have been tabulated with the counts in ascending order and the variate values $x = \frac{1}{2}[\log r + \log(r+1)]$ (i.e. the mid-points of the frequency intervals), obtained from Table 3 are placed alongside. The rest of the calculations follow exactly the order explained in § 2. A comparison of observed with expected frequencies is made, and the values of χ^2 obtained indicate no significant departures from hypothesis in the two samples.

The author wishes to thank Dr T. R. Vernon and Mr K. P. Lamb, of Plant Diseases Division, Auckland, for permission to use their data on spore counts and aphid counts respectively.

REFERENCES

Bliss, C. I. (1937). The calculation of the time-mortality curve. *Ann. Appl. Biol.* **24**, 815.

Gjeddebaek, N. F. (1949). Contributions to the study of grouped observations. Application of the method of maximum likelihood in case of normally distributed observations. *Skand. AktuarTidskr.* **32**, 135.

Hald, A. (1949). Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point. *Skand. AktuarTidskr.* **32**, 119.

Hartley, H. O. (1950). A simplified form of Sheppard's correction formulae. *Biometrika*, **37**, 145.

Spiller, D. (1948). Truncated log-normal and root-normal frequency distributions of insect populations. *Nature, Lond.*, **162**, 530.

Upholt, W. M. & Craig, R. (1940). A note on the frequency distribution of black scale insects. *J. Econ. Ent.* **33**, 113.

Williams, C. B. (1940). A note on the statistical analysis of sentence length as a criterion of literary style. *Biometrika*, **31**, 356.

Yule, G. U. (1944). *Statistical Study of Literary Vocabulary.* Cambridge University Press.

A CHART FOR THE INCOMPLETE BETA-FUNCTION AND THE CUMULATIVE BINOMIAL DISTRIBUTION

By H. O. HARTLEY and E. R. FITCH

1. Purpose of the chart

The present chart gives values of the Incomplete Beta-Function

$$P = I_x(a,b) = \int_0^x \xi^{a-1}(1-\xi)^{b-1}\,d\xi \Big/ \int_0^1 \xi^{a-1}(1-\xi)^{b-1}\,d\xi \qquad (1)$$

to graphical accuracy. It thereby provides, in compact form, approximate answers to the numerous problems in which this function arises. Of these we only mention the following two examples.

(1·1) *The cumulative binomial distribution*

If in repeated, independent trials the expected proportion of successful trials is p, find the chance P of observing at least a successful trials in a total of n trials. This chance is given by

$$P = \sum_{i=a}^{n} \binom{n}{i} p^i(1-p)^{n-i} = I_p(a, n-a+1). \qquad (2)$$

(1·2) *The distribution of the variance ratio F*

If s_1^2 and s_2^2 are two independent sample variances based on ν_1 and ν_2 degrees of freedom, find the chance that the variance ratio $F = s_1^2/s_2^2$ exceeds a given level F_0. This chance is given by

$$P(F \geqslant F_0) = I_{x_0}(\tfrac{1}{2}\nu_2, \tfrac{1}{2}\nu_1), \quad \text{where} \quad x_0 = \nu_2/(\nu_2 + \nu_1 F_0). \qquad (3)$$

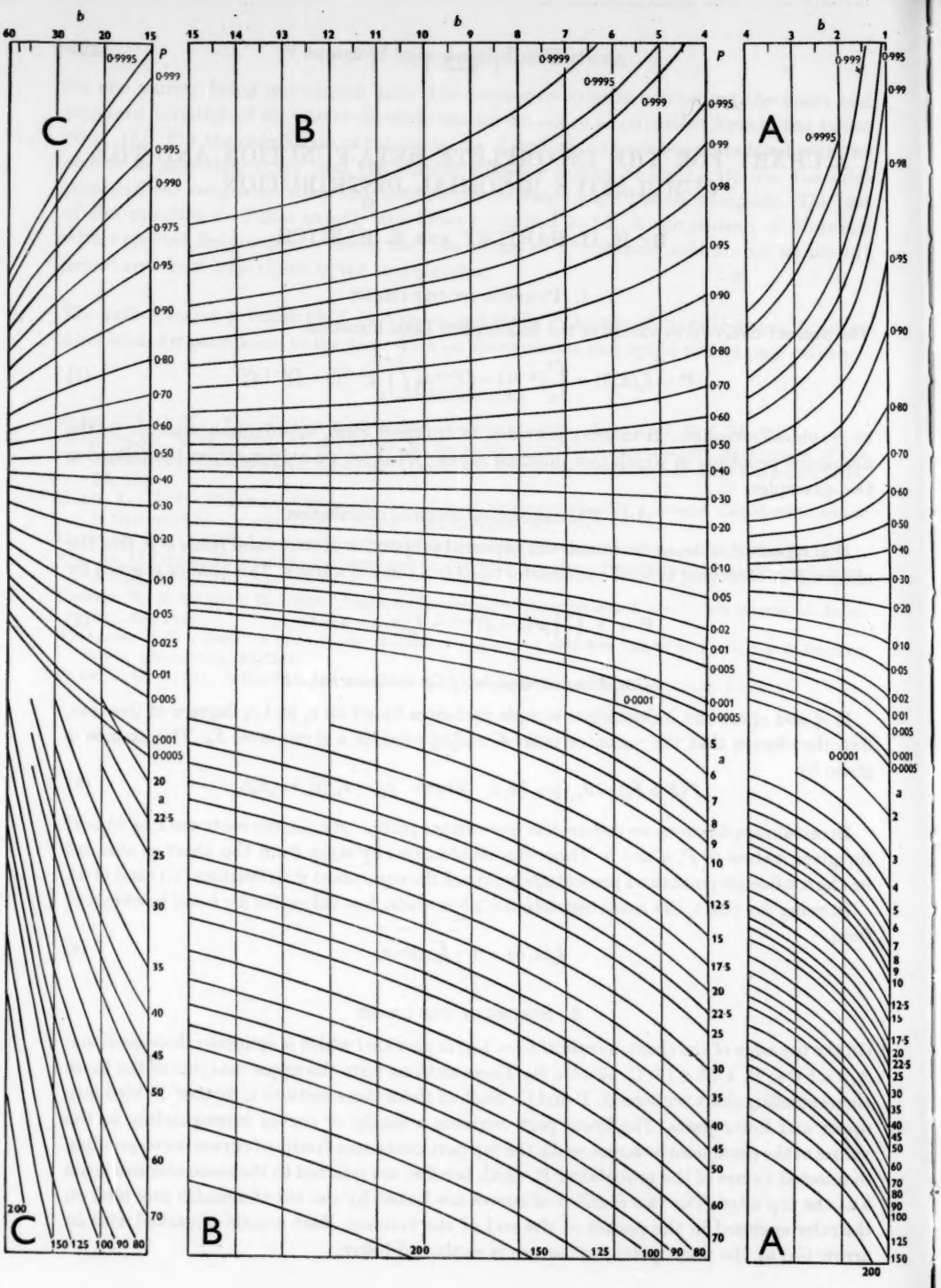
In certain applications we require the 'percentage points' of x, i.e. the roots x of $I_x(a,b) = P$, for given values of P, a and b. These can be obtained by sight from the chart as also can be the confidence points and percentage points of the equivalent distributions (1·1) and (1·2).

In using the chart, the usual restriction $a \geqslant b$ is made, but values for $a < b$ can be obtained from

$$I_x(a,b) = 1 - I_{1-x}(b,a). \qquad (4)$$

2. Scales of the chart

On the top edge of the chart, a scale of b ($= \frac{1}{2}\nu_1$) is provided which is split into three sections: A, $1 \leqslant b \leqslant 4$; B, $4 \leqslant b \leqslant 15$; C, $15 \leqslant b \leqslant 60$. These sections form the upper margins of the three corresponding chart sections A, B and C. Each of these three sections is further divided into upper and lower parts. The lower part contains a family of curves corresponding to the values of the parameter a shown, while the top part contains a family of curves corresponding to selected values of the probability P. Both families are referred to the same abscissa scale b at the top edge. The two families of curves are linked by the aid of x-scales provided on the ruler enclosed in the pocket at the end of the volume. Each x-scale is marked with an arrow (→) at the pivot point, whose use is explained below.


C
B
A
b
P
a
0·9995
0·9999
0·999
0·0001
200
150
125
100
90
80

3. USE OF THE CHART

(3·1) *To find $P = I_x(a, b)$ for given a, b and x*

Choosing the appropriate scale (A, B or C) place the ruler vertically on the chart, so that its x-scale intersects the abscissa scale (at top of chart) at the given value of b. Slide the ruler up or down this vertical until its pivot mark (→) is at the intersection of the vertical with the curve corresponding to the given value of a in the bottom part. With the x-scale in this position, the required value of $P = I_x(a, b)$ can be read off in the top part opposite the given x.

Examples. (1) Given $a = 30$, $b = 5$, it is required to find $I_x(a, b)$ for $x = 0{\cdot}70$, $0{\cdot}94$ and $0{\cdot}96$. Using section B we find:

x	From chart	For comparison, from tables of $I_x(a, b)$*
0·70	0·012	0·012
0·94	0·95(2)	0·949
0·96	0·99(0)	0·989

* *Tables of the Incomplete Beta-Function*, edited by K. Pearson (1934).

(2) In binomial theory (see (1·1) above), given that $p = 0{\cdot}2$ and $n = 25$, it is required to find the chance of obtaining, at least, $a = 10$ successes. This chance is given by $I_{0{\cdot}2}(10, 16)$ or $1 - I_{0{\cdot}8}(16, 10)$, so that, using section B, we find it to be equal to $1 - 0{\cdot}985 = 0{\cdot}015$.

(3·2) *To find the percentage points of the Incomplete Beta-Function*

(1) Given $a = 4$ and $b = 2$, it is required to find the upper and lower 10 % points. Using section A we find:

P	x	For comparison, from tables $I_x(a, b)$
0·90	0·89	0·888
0·10	0·42	0·416

(3·3) *To find the probability integral of F and to evaluate the percentage points of the distribution*

Using the relation (3), we proceed as in (3·1), putting $a = \frac{1}{2}\nu_2$, $b = \frac{1}{2}\nu_1$ and x as given by (3).

Example. Given $\nu_1 = 40$, $\nu_2 = 120$, find the chance that $F \geqslant 1{\cdot}5$. We have

$$a = 60, \quad b = 20, \quad x = 120/180 = 0{\cdot}667,$$

and hence, using section C, we find $P\{F \geqslant 1{\cdot}5\} = 0{\cdot}05(4)$. Merrington & Thompson's (1943) tables of the percentage points give $F = 1{\cdot}495$ for the upper 5 % point of F.

4. THEORETICAL BASIS OF THE CHART

The Incomplete Beta-Function involves the *four* parameters a, b, x and P.

In K. Pearson's (1934) tables P is tabulated as a function of x and a (called p) in a *series* of double-entry tables progressing by values of the fourth parameter b (called q). In con-

sidering any other form of tabulation or graphical representation of this four-parameter relationship it would, at first sight, appear that a *series* of tables or charts would be required. The reduction of the data to a *single* sheet is here achieved with the aid of the approximate formula

$$-\log(-\log x) = f(P,b)+g(a,b), \tag{5}$$

where

$$f(P,b) = \log\{\nu_1/\chi^2(P,2b)\},$$

$$g(a,b) = \log\{\tfrac{1}{2}(2a-1)/b+\tfrac{1}{2}\}-\log\{1+\tfrac{1}{8}(b-1)(2b+1)/(2a+b-1)^2\},$$

and $\chi^2(P,2b)$ is the $100P\,\%$ point of χ^2 for $2b$ degrees of freedom. Thus $\log(-\log x)$ is represented as a sum of two functions each involving only *two* parameters, and the chart is an obvious representation of this relationship.

Formula (5) was derived from an approximation recently published by Wise (1950).

REFERENCES

MERRINGTON, MAXINE & THOMPSON, CATHERINE M. (1943). *Biometrika*, **33**, 73.
PEARSON, K. (1934). *Tables of the Incomplete Beta-Function*. London: Biometrika Office.
WISE, E. (1950). *Biometrika*, **37**, 208.

THE EXPECTED FREQUENCIES IN A SAMPLE OF AN ANIMAL POPULATION IN WHICH THE ABUNDANCES OF SPECIES ARE LOG-NORMALLY DISTRIBUTED. PART I

BY P. M. GRUNDY

Rothamsted Experimental Station

INTRODUCTION

The sampling of animal populations by traps or other methods of random collection leads to interesting statistical problems, one of which is the subject of this paper. It is assumed that the counts from different species are independent Poisson variates, as occurs when all individuals have small and independent chances of capture. Accordingly, there is a probability

$$e^{-m} m^r/r!$$

of catching r individuals of any given species, m being the expected number. Here m may be called the *abundance* of the species concerned, but it must be recognized that this means abundance relative to both the intensity and the method of sampling. As in the paper of Fisher, Corbet & Williams (1943), the structure of the population may be investigated by studying the numbers of species represented in the sample by various numbers of individuals. The expected number of species with r individuals in the sample is

$$n_r = S(e^{-m} m^r/r!), \tag{1}$$

where S denotes summation over all species in the population.

In practice the number of species is large, and for theoretical purposes the frequency distribution of their abundances is taken to be continuous. Three such distributions which have been suggested are: the Eulerian or Type III distribution

$$\alpha\, e^{-\beta m} m^{\epsilon-1}\, dm,$$

the limit of this as $\epsilon \to 0$, and the log-normal distribution. The first of these alternatives leads to a negative binomial series of n_r, and the second to the logarithmic series investigated by Fisher *et al.* (1943); on the properties of these and related distributions, cf. Anscombe (1950). Observational evidence in favour of the normal distribution of $\log m$ has been adduced by Preston (1948); in the same paper he has worked out some examples of expected sample frequencies, using a discrete approximation to the normal distribution. His results indicate that the expected sample frequencies, grouped along the scale of r by octaves, conform well to the population curve after the first few octaves.

The object of this paper is to give tables and formulae for calculation of the n_r when m has a given log-normal distribution. Information of the kind thus provided, especially regarding the frequencies of rare species, is becoming increasingly necessary for the comparison of theory with the observed results as these accumulate. It is convenient to work with the proportions

$$\phi_r = n_r/N, \tag{2}$$

where N denotes the population number of species,

$$N = n_0 + n_1 + n_2 + \ldots.$$

In Part I, Tables 1 and 2 give the expected values of the proportion of species occurring in the sample, $1-\phi_0$, and of the proportion of singletons, ϕ_1. The other ϕ_r can be calculated, over the range of a and σ^2 anticipated, by an approximate method which will be described in Part II; it is only for ϕ_0 and ϕ_1 that the labour of constructing explicit tables appears to be justified. An approximate formula relevant in the fitting of the distribution to observational data is also in sight, but this cannot be discussed at the present stage.

The notation used for the parameters of the distribution of abundance is as follows. The variance of the *natural* logarithm $\ln m$ is denoted by σ^2; the distribution of $\ln m$ is assumed to be normal with mean $\ln a$, so that a denotes the median abundance. Thus for all $r \geqslant 0$

$$\phi_r = \int_0^\infty \frac{e^{-m} m^r}{r!} f(m)\, dm, \tag{3}$$

where

$$f(m) = \frac{1}{m\sigma\sqrt{(2\pi)}} \exp\left\{\frac{-1}{2\sigma^2}\left(\ln\frac{m}{a}\right)^2\right\}. \tag{4}$$

The expected number of individuals in the sample (the integral of $Nmf(m)\, dm$, which reduces to a formula involving the normal integral) is $Na\, e^{\frac{1}{2}\sigma^2}$.

The Jahnke & Emde tables (1945) include a table, due to Wagner (1913), of the 'after-effect function' $(\sigma\sqrt{(2\pi)}/a)\,\phi_1$ as a function of a and $1/\sigma\sqrt{2}$. However, both the range and the irregular spacing of the values of a make that table inconvenient for application to the present problem. A fresh table of ϕ_1 has therefore been calculated (Table 2), as well as the table of $1-\phi_0$ (Table 1), proceeding at equal intervals of $\log a = \log_{10} a$. The methods used in the calculation and checking of these tables are discussed below.

Interpolation

The probits (i.e. normal equivalent deviates + 5) of the entries in Table 1 have fairly small and slowly varying second differences with respect to $\log a$, and it is advantageous to use the probits in interpolation with respect to $\log a$. A similar advantage is obtained by taking the probits of the differences between corresponding values in Tables 1 and 2 (i.e. by using the probit of $1-\phi_0-\phi_1$) in interpolation with respect to $\log a$. The transformation to probits is suggested by Preston's approach (1948) (cf. Fig. 1 of the present paper). However, the transformation is not advisable in interpolation with respect to σ^2.

Probits of Table 1

Preston (1948) has introduced the concept of the 'veil line' separating the less abundant species ($m < 1$) from those more abundant ($m > 1$). Heuristically, the sample may be supposed to include most of the species to the right of the veil line, and few of those to the left. This line of thought suggests examination of the probits of the entries in Table 1, and leads to the expression*

$$5 + \frac{\ln a}{\sigma} \tag{5}$$

as a possible approximation to the probit of $1-\phi_0$. In this context it was suggested to the author by Dr C. B. Williams that the level of abundance $m = \frac{1}{2}$ might give a better position for the veil line. This corresponds to replacing $\ln a$ by $\ln 2a$ in the above formula, but it has

* Preston does not, however, seem to use this, since he adopts a method of grouping into octaves by which half the species appearing as singletons in the sample are assigned to the left of the veil line.

Table 1. *Expected sample number of species as a fraction of N, i.e.* $1-\phi_0$

(Population parameters: N = number of species; a = median and σ^2 = variance of natural logarithms of the abundances of species.)

$\log_{10} a$	$\sigma^2=2$	$\sigma^2=3$	$\sigma^2=4$	$\sigma^2=5$	$\sigma^2=6$	$\sigma^2=7$	$\sigma^2=8$	$\sigma^2=9$	$\sigma^2=10$	$\sigma^2=11$	$\sigma^2=12$	$\sigma^2=13$	$\sigma^2=14$	$\sigma^2=15$	$\sigma^2=16$
−2·0	0·0251	0·0363	0·0488	0·0616	0·0742	0·0863	0·0978	0·1087	0·1189	0·1285	0·1374	0·1459	0·1538	0·1613	0·1683
−1·75	0·0425	0·0588	0·0753	0·0912	0·1061	0·1199	0·1326	0·1443	0·1550	0·1649	0·1741	0·1826	0·1904	0·1978	0·2047
−1·5	0·0703	0·0922	0·1126	0·1311	0·1475	0·1622	0·1754	0·1872	0·1979	0·2076	0·2165	0·2246	0·2321	0·2390	0·2455
−1·25	0·1129	0·1397	0·1628	0·1824	0·1992	0·2137	0·2264	0·2376	0·2476	0·2565	0·2645	0·2718	0·2785	0·2847	0·2903
−1·0	0·1743	0·2037	0·2269	0·2456	0·2611	0·2741	0·2852	0·2949	0·3034	0·3109	0·3176	0·3237	0·3292	0·3342	0·3388
−0·75	0·2574	0·2846	0·3046	0·3200	0·3323	0·3424	0·3509	0·3581	0·3644	0·3700	0·3749	0·3793	0·3832	0·3868	0·3901
−0·5	0·3613	0·3805	0·3936	0·4032	0·4107	0·4167	0·4217	0·4258	0·4294	0·4325	0·4352	0·4376	0·4398	0·4418	0·4436
−0·25	0·4806	0·4865	0·4899	0·4920	0·4935	0·4946	0·4954	0·4960	0·4965	0·4969	0·4972	0·4975	0·4978	0·4980	0·4981
0	0·6050	0·5953	0·5878	0·5819	0·5771	0·5731	0·5696	0·5666	0·5639	0·5616	0·5595	0·5576	0·5558	0·5543	0·5528
0·25	0·7221	0·6986	0·6815	0·6684	0·6578	0·6491	0·6417	0·6353	0·6297	0·6248	0·6204	0·6164	0·6128	0·6095	0·6065
0·5	0·8208	0·7890	0·7655	0·7472	0·7323	0·7198	0·7092	0·7001	0·6920	0·6849	0·6785	0·6728	0·6676	0·6628	0·6584
0·75	0·8951	0·8620	0·8361	0·8152	0·7978	0·7831	0·7704	0·7592	0·7494	0·7406	0·7327	0·7255	0·7190	0·7130	0·7075
1·0	0·9447	0·9159	0·8916	0·8708	0·8529	0·8374	0·8237	0·8115	0·8007	0·7908	0·7819	0·7738	0·7663	0·7594	0·7530
1·25	0·9739	0·9525	0·9323	0·9137	0·8971	0·8821	0·8685	0·8563	0·8451	0·8349	0·8255	0·8169	0·8089	0·8014	0·7945
1·5	0·9890	0·9752	0·9601	0·9451	0·9308	0·9174	0·9049	0·8933	0·8825	0·8725	0·8632	0·8545	0·8464	0·8388	0·8316
1·75	0·9959	0·9881	0·9780	0·9668	0·9554	0·9442	0·9333	0·9229	0·9131	0·9038	0·8949	0·8866	0·8787	0·8712	0·8641
2·0	0·9987	0·9947	0·9886	0·9809	0·9724	0·9636	0·9547	0·9459	0·9373	0·9290	0·9210	0·9133	0·9059	0·8988	0·8921
2·25	0·9996	0·9979	0·9944	0·9896	0·9837	0·9771	0·9702	0·9631	0·9560	0·9488	0·9419	0·9350	0·9284	0·9219	0·9156
2·5	0·9999	0·9992	0·9975	0·9946	0·9908	0·9862	0·9811	0·9756	0·9699	0·9640	0·9581	0·9523	0·9465	0·9407	0·9351
2·75	1·0000	0·9997	0·9989	0·9974	0·9950	0·9920	0·9884	0·9843	0·9799	0·9753	0·9705	0·9657	0·9608	0·9558	0·9509
3·0	1·0000	0·9999	0·9996	0·9988	0·9974	0·9955	0·9931	0·9902	0·9870	0·9835	0·9797	0·9758	0·9718	0·9677	0·9635

Table 2. *Expected number of species appearing as singletons in the sample, as a fraction of N, i.e.* ϕ_1

(Population parameters as in Table 1.)

$\log_{10} a$	$\sigma^2=2$	$\sigma^2=3$	$\sigma^2=4$	$\sigma^2=5$	$\sigma^2=6$	$\sigma^2=7$	$\sigma^2=8$	$\sigma^2=9$	$\sigma^2=10$	$\sigma^2=11$	$\sigma^2=12$	$\sigma^2=13$	$\sigma^2=14$	$\sigma^2=15$	$\sigma^2=16$
−2·0	0·0234	0·0312	0·0381	0·0436	0·0480	0·0513	0·0538	0·0556	0·0570	0·0579	0·0586	0·0590	0·0592	0·0593	0·0593
−1·75	0·0381	0·0476	0·0548	0·0599	0·0634	0·0658	0·0673	0·0682	0·0686	0·0688	0·0687	0·0684	0·0681	0·0676	0·0671
−1·5	0·0598	0·0695	0·0755	0·0789	0·0807	0·0814	0·0815	0·0811	0·0805	0·0796	0·0787	0·0777	0·0766	0·0756	0·0745
−1·25	0·0892	0·0963	0·0991	0·0995	0·0988	0·0973	0·0956	0·0937	0·0918	0·0899	0·0880	0·0862	0·0845	0·0828	0·0812
−1·0	0·1251	0·1260	0·1235	0·1199	0·1160	0·1122	0·1085	0·1051	0·1019	0·0989	0·0961	0·0936	0·0912	0·0890	0·0869
−0·75	0·1633	0·1546	0·1457	0·1377	0·1307	0·1245	0·1190	0·1142	0·1099	0·1060	0·1025	0·0993	0·0964	0·0938	0·0913
−0·5	0·1962	0·1772	0·1624	0·1506	0·1410	0·1330	0·1262	0·1204	0·1152	0·1107	0·1067	0·1031	0·0998	0·0968	0·0941
−0·25	0·2151	0·1889	0·1704	0·1566	0·1456	0·1367	0·1292	0·1229	0·1174	0·1126	0·1084	0·1046	0·1011	0·0980	0·0952
0	0·2134	0·1866	0·1682	0·1545	0·1438	0·1351	0·1278	0·1216	0·1162	0·1115	0·1074	0·1037	0·1003	0·0972	0·0945
0·25	0·1901	0·1701	0·1557	0·1446	0·1357	0·1282	0·1219	0·1165	0·1118	0·1076	0·1038	0·1004	0·0974	0·0946	0·0920
0·5	0·1512	0·1428	0·1351	0·1282	0·1222	0·1169	0·1123	0·1081	0·1044	0·1010	0·0979	0·0951	0·0925	0·0901	0·0879
0·75	0·1068	0·1102	0·1096	0·1076	0·1051	0·1024	0·0997	0·0971	0·0946	0·0923	0·0900	0·0879	0·0860	0·0841	0·0824
1·0	0·0667	0·0779	0·0831	0·0855	0·0862	0·0860	0·0854	0·0844	0·0833	0·0820	0·0808	0·0795	0·0782	0·0770	0·0758
1·25	0·0367	0·0503	0·0588	0·0641	0·0674	0·0693	0·0704	0·0710	0·0711	0·0710	0·0706	0·0702	0·0696	0·0690	0·0683
1·5	0·0177	0·0297	0·0388	0·0455	0·0502	0·0536	0·0560	0·0578	0·0589	0·0597	0·0602	0·0605	0·0606	0·0606	0·0605
1·75	0·0075	0·0160	0·0239	0·0304	0·0356	0·0397	0·0430	0·0455	0·0474	0·0489	0·0501	0·0510	0·0516	0·0521	0·0525
2·0	0·0028	0·0079	0·0137	0·0192	0·0241	0·0282	0·0317	0·0346	0·0370	0·0390	0·0406	0·0419	0·0430	0·0439	0·0447
2·25	0·0009	0·0035	0·0073	0·0114	0·0155	0·0192	0·0226	0·0255	0·0280	0·0302	0·0321	0·0337	0·0351	0·0363	0·0373
2·5	0·0002	0·0014	0·0036	0·0064	0·0095	0·0125	0·0154	0·0181	0·0206	0·0228	0·0247	0·0264	0·0280	0·0293	0·0305
2·75	0·0001	0·0005	0·0017	0·0034	0·0055	0·0078	0·0102	0·0125	0·0146	0·0167	0·0185	0·0203	0·0218	0·0232	0·0245
3·0	0·0000	0·0002	0·0007	0·0017	0·0031	0·0047	0·0065	0·0083	0·0101	0·0119	0·0136	0·0152	0·0166	0·0180	0·0193

been found that some adjustment of the variance is also desirable. Over the range of Table 1 the empirical formula

$$4{\cdot}97 + \frac{\ln 2a}{\sqrt{(\sigma^2 + 1{\cdot}4)}} \tag{6}$$

gives about as good an approximation to the probit of $1 - \phi_0$ as is possible for a function linear in $\ln a$. There are also theoretical reasons for using an approximation of this type, which the author hopes to discuss in a future paper. A comparison of the correct values (based on Table 1) with the expressions (5) and (6) is shown in Fig. 1, and it will be seen that for the rather small variance $\sigma^2 = 2$ the graph of the probits shows appreciable curvature; when $\sigma^2 = 5$ the curvature, though perceptible, is very slight.

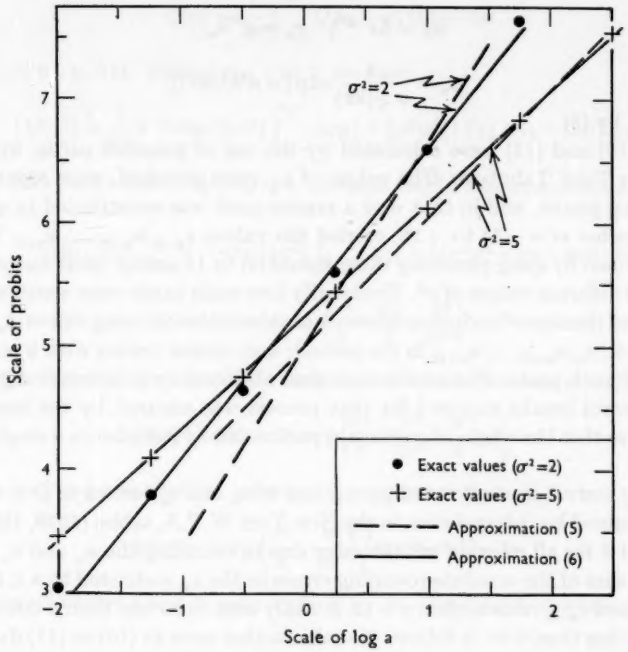


Fig. 1. Proportion of species expected in the sample. Comparison of correct values (Table 1) with approximate formulae, on probit scale.

METHOD OF CALCULATING TABLES 1 AND 2

The substitution $x = \ln m$ throws the integral (3) into the form

$$\phi_r = \int_{-\infty}^{\infty} \frac{1}{r!\,\sigma\sqrt{(2\pi)}} \exp\left\{-e^x + rx - \frac{1}{2\sigma^2}(x - \ln a)^2\right\} dx. \tag{7}$$

In the cases $r = 0$ and 1, this integral was evaluated by the approximate formula

$$\int_{-\infty}^{\infty} g(x)\,dx \doteqdot h \sum_{-\infty}^{\infty} g(nh) \quad (h > 0). \tag{8}$$

The error involved in (8) is discussed below, and it is shown that in the application to (7) over the range of Tables 1 and 2 the error is negligible when

$$h = \tfrac{1}{2}\ln 10 = 0{\cdot}575646273. \tag{9}$$

Accordingly, to the required accuracy, ϕ_0 is given by

$$\sum_{-\infty}^{\infty} u_n z_{n-k}, \tag{10}$$

and ϕ_1 by

$$\sum_{-\infty}^{\infty} v_n z_{n-k}, \tag{11}$$

where

$$k = h^{-1}\ln a = 4\log a,$$

$$u_n = h\,e^{-e^{nh}}, \quad v_n = e^{nh}u_n,$$

$$z_n = \frac{1}{\sigma\sqrt{(2\pi)}}\exp[-n^2h^2/2\sigma^2],$$

and h is given by (9).

The sums (10) and (11) were calculated by the use of punched cards, with a Hollerith Junior Rolling Total Tabulator. The values of u_n, once punched, were repeatedly fed into the reproducing punch, and in that way a master pack was constructed in which the card with serial number $n(=-34$ to $+22)$ carried the values $u_{n-8}, u_{n-7}, \ldots, u_{n+2}$. The first main pack was obtained by gang-punching these values on to 15 sets of cards carrying the values of z_n for the 15 different values of σ^2. Eventually four main packs were constructed, in which the cards carried (besides σ^2 and other reference numbers) the following values: $z_n, u_{n-8}, \ldots, u_{n+2}$ in the first pack; $z_n, u_{n+2}, \ldots, u_{n+12}$ in the second; and similar values with u replaced by v in the third and fourth packs. The results were then obtained by progressive digiting. The full number of control-breaks required for that process was ensured by the insertion of subsidiary packs, so that the whole of each main pack could be included in a single run through the tabulator.

Values of z_n correct to six decimal places, and of u_n and v_n correct to five, were used, the data being obtained by interpolation in the New York W.P.A. tables (1939, 1942). Since Σz_n was less than 1·8 for all relevant σ^2, the error due to rounding the u_n and v_n was less than 0·000009. The sum of the absolute rounding errors in the z_n, restricted to $n \leqslant 12$ because the rounded u_{n+k} and v_{n+k} vanish when $n > 12$, is easily seen to be less than 0·000025. Since the u_n and v_n were less than 0·58, it follows that the further error in (10) or (11) due to rounding of the z_n was less than 0·000015. The actual rounding errors (previous to rounding of the final results to four decimal places) were probably less than half as much.

The remaining source of error, the approximation (8), has already been mentioned. This error (for any $h > 0$) is in general given by *Poisson's summation formula*, namely,

$$\int_{-\infty}^{\infty} g(x)\,dx = h\sum_{-\infty}^{\infty} g(nh) - \sum_{n=1}^{\infty}\left\{Q\left(\frac{2\pi n}{h}\right) + Q\left(\frac{-2\pi n}{h}\right)\right\}, \tag{12}$$

where

$$Q(y) = \int_{-\infty}^{\infty} e^{ixy}\,g(x)\,dx. \tag{13}$$

Sufficient conditions for the validity of (12) are (Titchmarsh, 1948, Theorem 45) that, in the whole range $-\infty < x < \infty$, $g(x)$ be continuous and expressible as the difference between two bounded non-decreasing functions, and that the integral on the left-hand side of (12) converge.

We take $g(x)$ to be the integrand of (7), and note that it is the product of

$$g_1(x) = \exp(-e^x + \tfrac{1}{2}x),$$

and

$$g_2(x) = \frac{1}{r!\,\sigma\sqrt{(2\pi)}} \exp\left\{(r-\tfrac{1}{2})x - \frac{1}{2\sigma^2}(x-\ln a)^2\right\}.$$

To these functions correspond the known transforms

$$Q_1(y) = \Gamma(\tfrac{1}{2}+iy),$$

$$Q_2(y) = \frac{1}{r!}\, a^{r-\frac{1}{2}+iy} \exp[\tfrac{1}{2}\sigma^2(r-\tfrac{1}{2}+iy)^2].$$

Now

$$Q(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Q_1(y-w)\, Q_2(w)\, dw$$

(Titchmarsh, 1948, p. 51). Taking $r = 0$ or 1, we have

$$|\,Q(y)\,| \leqslant \frac{1}{2\pi}\, a^{r-\frac{1}{2}} \exp[\tfrac{1}{8}\sigma^2] \int_{-\infty}^{\infty} \exp[-\tfrac{1}{2}\sigma^2 w^2]\,|\,\Gamma(\tfrac{1}{2}+iy-iw)\,|\,dw.$$

It is shown in text-books that $|\,\Gamma(\tfrac{1}{2}+iw)\,| = \sqrt{(\pi \operatorname{sech} \pi w)}$ for real w, and since $\operatorname{sech} \pi w < 2e^{-\pi w}$ it follows that

$$|\,Q(y)\,| < \frac{1}{\sqrt{(2\pi)}}\, a^{r-\frac{1}{2}} \exp[\tfrac{1}{8}\sigma^2] \int_{-\infty}^{\infty} \exp[-\tfrac{1}{2}(\sigma^2 w^2 + \pi y - \pi w)]\, dw$$

$$= a^{r-\frac{1}{2}}\sigma^{-1} \exp\left\{\frac{\sigma^2}{8} - \frac{\pi y}{2} + \frac{\pi^2}{8\sigma^2}\right\}.$$

A similar inequality with the sign of y changed is obtained from $\operatorname{sech} \pi w < 2\,e^{\pi w}$, and thus

$$|\,Q(y)\,| < a^{r-\frac{1}{2}}\sigma^{-1} \exp\left\{\frac{\sigma^2}{8} + \frac{\pi^2}{8\sigma^2} - \frac{\pi}{2}\,|\,y\,|\right\}.$$

Substitution in (12) shows that the error in (8), applied to ϕ_0 or ϕ_1, is less than

$$2a^{r-\frac{1}{2}}\sigma^{-1} \exp\left(\frac{\sigma^2}{8} + \frac{\pi^2}{8\sigma^2}\right) \Big/ (\exp[\pi^2/h] - 1).$$

Over the range of Tables 1 and 2, with h given by (9), this quantity is less than 0·000005.

CHECKS

The values of z were checked by summation, using a known result of applying (12) to the normal curve, that Σz_n differs from $1/h$ by nearly $2h^{-1}\exp[-2\pi^2\sigma^2/h^2]$. The values of u and v and the original punching, the reproducing, gang-punching and sorting, were checked by methods which do not call for special comment. The Hollerith results for different columns of z were combined on desk machines, and a final check of that process and of the Hollerith work was obtained by calculating independently the quantities

$$\sum_n z_n \left(\sum_{k=-8}^{+12} u_{n+k}\right), \quad \sum_n z_n \left(\sum_{k=-8}^{+12} v_{n+k}\right)$$

for all σ^2. This brought to light the fact that one of several cards damaged in the sorter had been incorrectly renewed, but eventually exact agreement was obtained.

SUMMARY

The expected number of species having r representatives in a sample of an animal population is considered. It is assumed that the population distribution of abundance is log-normal, and that the sampling process results in Poisson variation of the number obtained from each species. Tables dealing with the cases $r = 0, 1$ are provided, and their construction, checking and accuracy discussed.

The author is indebted to Dr C. B. Williams for suggesting this investigation, to Dr F. Yates for his interest and encouragement, and to several colleagues for advice on the Hollerith work, especially Mr G. V. Dyke. Thanks are also due to A. C. Sphinx Sparking Plugs Ltd for the use of their Reproducing Punch.

REFERENCES

ANSCOMBE, F. J. (1950). Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, **37**, 358–82.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58.

JAHNKE, E. & EMDE, F. (1945). *Tables of Functions with Formulae and Curves*. New York.

NEW YORK W.P.A. (1939). *Tables of the Exponential Function* e^x.

NEW YORK W.P.A. (1942). *Tables of Probability Functions*, **2**.

PRESTON, F. W. (1948). The commonness, and rarity, of species. *Ecology*, **29**, 254–83.

TITCHMARSH, E. C. (1948). *Introduction to the Theory of Fourier Integrals*, 2nd ed. Oxford.

WAGNER, K. W. (1913). *Elektrotech. Z.* **34**, 1279–81.

TABLES OF SYMMETRIC FUNCTIONS. PARTS II AND III

By F. N. DAVID and M. G. KENDALL

1. In a previous paper (1949) we presented tables of the augmented monomial symmetric functions in terms of the one-part symmetric functions and vice versa. The present paper presents the monomials in terms of unitary functions (and vice versa) and the unitary functions in terms of homogeneous product-sums. It is convenient to recall the definitions of these functions:

(*a*) The monomial symmetrics (which we denote for short by M) are typified by

$$(p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}) = \Sigma(x_i^{p_1} x_j^{p_1} \dots x_q^{p_2} x_r^{p_2} \dots x_u^{p_s} x_v^{p_s}), \tag{1}$$

where the summation takes place over all suffixes $i, j, \dots, v$ which are different. The augmented symmetrics are defined by

$$[p_1^{\pi_1} \dots p_s^{\pi_s}] = (p_1^{\pi_1} \dots p_s^{\pi_s}) \pi_1! \dots \pi_s!. \tag{2}$$

(*b*) The unitary symmetrics are defined by

$$a_r = (1^r) = \Sigma x_i x_j \dots x_v \quad (i \neq j \neq \dots \neq v). \tag{3}$$

We denote them for short by U.

(*c*) The one-part symmetrics or power-sums are defined by

$$s_r = (r) = \Sigma x_i^r. \tag{4}$$

We denote them by S.

(*d*) The homogeneous product-sums are defined by

$$\frac{1}{T} = 1 + h_1 t + \dots + h_r t^r + \dots, \tag{5}$$

where

$$T = (1 - x_1 t)(1 - x_2 t) \dots = 1 - a_1 t + a_2 t^2 - \dots. \tag{6}$$

We denote them by H.

Briefly, Part I of these tables presented MS-SM tables. The present set gives MU-UM (Tables 2) and UH-HU (Tables 3). A third paper in preparation will give MH-HM, US-SH and SH-HS. In all cases we proceed as far as tables of weight 12.

2. When Part I of these tables was prepared we were unaware of some work of Kerawala & Hanafi (1941, 1942, 1948), who have given tables of MS (but not SM) for weights 9, 10, 11 and 12. Sukhatme (1938) had previously given tables up to $w = 8$. We have compared these tables with the corresponding parts of our own for $w = 2$ to 8 and $w = 12$; we have not been able to consult the tables for $w = 9$ to 11. There is complete agreement up to weight 8. For weight 12 we find 8 points of disagreement with the table by Kerawala & Hanafi, but a re-check of our own figures confirms them. There is therefore reason to hope that our own tables are free from error.

3. Tables 2 give U in terms of M by reading horizontally up to the main diagonal, e.g. from Table 2.11,

$$a_8 a_3 = 165(1^{11}) + 36(21^9) + 7(2^2 1^7) + (2^3 1^5), \tag{7}$$

M is given in terms of U by reading downwards to the main diagonal, e.g. from Table 2.11,

$$(2^3 1^5) = -77 a_{11} + 27 a_{10} a_1 - 7 a_9 a_2 + a_8 a_3. \tag{8}$$

It should be noticed that the order of the column headings in Tables 2 is not the same as that of the rows. In Tables 1 we introduced the augmented symmetric functions mainly in order to produce units in the main diagonal and hence to incorporate both MS and SM in one square array. This obviously saves a lot of space, and we have achieved the same result in Tables 2 by a special order of the columns. The rows, like the rows and columns of Table 1, are in 'alphabetical' order, that is to say, the partitions whose highest part is p all come before those whose highest part is less than p, and so on. In the columns of Tables 2 the order is by number of parts, those with more coming in front of those with fewer; of those with the same number, the later are those with unitary parts, then come those with parts 2 and so on.

4. Tables 3 give H in terms of U by reading horizontally right across the table, e.g. from Table 3.5,

$$h_4h_1 = -a_4a_1 + 2a_3a_1^2 + a_2^2a_1 - 3a_2a_1^3 + a_1^5. \tag{9}$$

If we combine (5) and (6) we have

$$\frac{1}{\sum_{r=0}^{\infty} h_r t^r} = \sum_{r=0}^{\infty} a_r(-1)^r t^r, \tag{10}$$

which is equivalent to

$$\frac{1}{\Sigma a_r(-1)^r t^r} = \Sigma h_r t^r.$$

Substituting t for $-t$ we find

$$\frac{1}{\Sigma a_r t^r} = \Sigma h_r(-1)^r t^r, \tag{11}$$

which is exactly of the form (10) with a and h interchanged. It follows that any relation expressing h's in terms of a's is equally valid if we interchange h and a in it; and consequently the same tables express U in terms of H by reading horizontally right across the table, e.g. corresponding to (9), we have

$$a_4a_1 = -h_4h_1 + 2h_3h_1^2 + h_2^2h_1 - 3h_2h_1^3 + h_1^5. \tag{12}$$

For Tables 3, therefore, one triangular array covers both direct and inverse relations.

Uses of Tables 2 and 3

5. Tables 2 and 3 have various uses in the theories of statistics and probability. We exemplify a few of them.

(*a*) *Sampling from a finite population.* Isserlis (1931, p. 593) gave what is equivalent to MU tables up to weight 8, omitting terms in a_1 which he did not need in his discussion of sampling from a finite population. (He refers to tables up to weight 10 by Salmon, who took them from Cayley (1857).) Isserlis's method of evaluating sampling moments for a finite population of size N is to expand the statistic under investigation in terms of the M-functions and then calculate mean values by the simple relation

$$\mathscr{E}(p_1^{\pi_1} \dots p_s^{\pi_s})_n = \frac{n^{(\varpi)}}{N^{(\varpi)}}(p_1^{\pi_1} \dots p_s^{\pi_s})_N, \tag{13}$$

where the suffixes to the monomials denote respectively that they relate to a sample of n and a population of N, $n^{(\varpi)} = n(n-1)\dots(n-\varpi+1)$, and ϖ is the weight of the monomial, namely, $\Sigma\pi$. He then expresses the population M-functions in terms of U-functions and thence in terms of S-functions. By the use of Tables 1 we can proceed direct from M- to S-functions, a course followed by Sukhatme (1944).

(*b*) *Distribution problems.* MacMahon (1915) gives a number of results concerning the distribution of objects of given type into boxes of given type. For instance, let there be p_1 objects of one type, p_2 of a second, ..., p_s of an sth type and m boxes. Then the number of ways of distributing the objects such that there are q_j objects in the jth box and no type of object appears more than once in the same box and no box is empty is the coefficient of $(p_1 p_2 \dots p_s)$ in $a_{q_1} a_{q_2} \dots a_{q_m}$ multiplied by the numerical coefficient of $x^{p_1+p_2+\dots+p_s} a_{q_1} a_{q_2} \dots a_{q_m}$ in the development of $(a_1 x + a_2 x^2 + \dots)^{q_1+q_2+\dots+q_m}$. Such coefficients are obtainable from Tables 2. For instance, if we have four red, three white, three black balls and one green ball, eleven altogether, they can be allocated so that there are two balls in each of five boxes and one in a sixth, no pair being of the same colour, in 6×430 ways; 430 being from Table 2.11 the coefficient of (43^21) in $a_2^5 a_1$, and the 6 being the coefficient of $x^{11} a_2 a_1^5$ in the expansion of $(a_1 x + a_2 x^2 + \dots)^6$.

Again, the number of ways of placing units in the compartments of a lattice with m rows and l columns so that not more than one unit occurs in any one compartment, and such that there are $\mu_1, \mu_2, \dots, \mu_m$ units in the successive rows and $\lambda_1, \lambda_2, \dots, \lambda_l$ in successive columns, is the coefficient of $(\lambda_1 \lambda_2 \dots \lambda_l)$ in the expansion of $a_{\mu_1} a_{\mu_2} \dots a_{\mu_m}$. Such results have applications in the enumeration of experimental designs.

(*c*) *Multipartite numbers.* MacMahon (1915, p. 155) shows that the number of compositions of the multipartite number $(pqr\dots)$ is the coefficient of the symmetric function $(pqr\dots)$ in the development of the algebraic fraction

$$\frac{1}{2\{1-2(a_1-a_2+a_3-\dots)\}}.$$

The number of compositions can thus be obtained by expanding the fraction and adding the coefficients of M in each relevant U product.

CONSTRUCTION OF THE TABLES

6. The UM half of Tables 2 was obtained by building up the table for any given weight from those of lower weights. For example, from Table 2.9

$$a_7 a_2 = 36(1^9) + 7(21^7) + (2^21^5),$$

and hence

$$\begin{aligned} a_7 a_2 a_1 = (1)\{a_7 a_2\} &= 36\{10(1^{10}) + (21^8)\} + 7\{8(21^8) + 2(2^21^6) + 7(31^7)\} \\ &\quad + \{6(2^21^6) + 3(2^31^4) + (321^5)\} \\ &= 360(1^{10}) + 92(21^8) + 20(2^21^6) + 7(31^7) + 3(2^31^4) + (321^5). \end{aligned}$$

The MU half was obtained by inverting the MU relations.

Four checks were available:

(*a*) There is a law of symmetry which asserts that the coefficient of $a_{p_1} \dots a_{p_s}$ in $(q_1 \dots q_r)$ is the coefficient of $a_{q_1} \dots a_{q_r}$ in $(p_1 \dots p_s)$ and vice versa; hence each coefficient checks against an independently calculated coefficient except in a few symmetrical cases such as the coefficient of (2^4) in a_2^4.

(*b*) If we put each member entering into a symmetric function equal to unity and take w members, there are simple arithmetical relations for each row and column; e.g. for $w = 10$ we then have $a_{10} = 1$, $a_9 = a_1 = 10$, $a_8 = a_2 = 45$, $a_7 = a_3 = 120$, $a_6 = a_4 = 210$, $a_5 = 252$. Also $(1^{10}) = 1$, $(21^8) = 90$, $(2^21^6) = 1260$, $(31^7) = 360$, etc. Thus, in the row corresponding to $a_7 a_2 a_1$ in Table 2.10 we check by the identity

$$54000 = 360 + 92(90) + 20(1260) + 7(360) + 3(4200) + 5040.$$

Each row and column was checked in this way:

(*c*) the D-operator of Hammond referred to in the Introduction to Tables 1 was used in the manner there described to provide a further check;

(*d*) the tables were collated with those published by Cayley (1857) for weights 2–10, and Durfee (1882) for weight 12. Tables for weight 11 have been given by Faà de Bruno (1876), but we have not been able to find a copy of this work in this country.

7. Tables 3 were much easier to calculate. Again any table for given weight was built up from tables for lower weights. Since $a_1 = h_1$ any function with a unit part is obtainable immediately from a previously calculated value. For non-unitary parts we have, for example,

$$\begin{aligned} h_4 h_2 &= (-a_2 + a_1^2)(-a_4 + 2a_3 a_1 + a_2^2 - 3a_2 a_1^2 + a_1^4) \\ &= a_4 a_2 - a_4 a_1^2 - 2a_3 a_2 a_1 + 2a_3 a_1^3 - a_2^3 + 4a_2^2 a_1^2 - 4a_2 a_1^4 + a_1^6. \end{aligned}$$

Three checks were used:

(*a*) the sum of coefficients in any row is zero;

(*b*) the relations between a and h being identical we can check by mutual orthogonality; e.g. in Table 3.4,

$$\begin{aligned} h_3 h_1 &= a_3 a_1 - 2a_2 a_1^2 + a_1^4 \\ &= (h_3 h_1 - 2h_2 h_1^2 + h_1^4) - 2(-h_2 h_1^2 + h_1^4) + h_1^4 \\ &= h_3 h_1; \end{aligned}$$

(*c*) the D-operator also gives an independent check for each relationship.

Values for higher weights

8. For weights higher than 12 individual values can be derived by operational processes as exemplified in the Introduction to Tables 1. We assume an expansion of, say, a given M-function of weight w in terms of the U-functions and operate by Hammond operators of order $(w-12)$, identifying the result with values given by Table 2.12. This will determine nearly all the coefficients and any remaining can be determined by operating by a higher-order operator and comparing with a Table 2 for a lower weight.

Full MU tables for weight 13 have been given by P. A. MacMahon (1882). It is thought that tables of weight 14 have been computed but not published. Tables of the 15th degree have been calculated by J. B. Shaw and printed as one of the publications of the Carnegie Institute of Washington.

We are indebted to Prof. D. Lehmer for some of the references.

REFERENCES

Cayley, A. (1857). *Philos. Trans.* **147**, 489.
David, F. N. & Kendall, M. G. (1949). *Biometrika*, **36**, 431.
Durfee, W. P. (1882). *Amer. J. Math.* **5** (5), 46.
Faà de Bruno (1876). *Théorie des Formes Binaires.* Brero, Turin.
Isserlis, L. (1931). *Proc. Roy. Soc.* A, **132**, 586.
Kerawala, S. M. & Hanafi, A. R. (1941). *Proc. Nat. Acad. Sci. India*, **11** (3), 51.
Kerawala, S. M. & Hanafi, A. R. (1942). *Proc. Nat. Acad. Sci. India*, **12** (2), 81.
Kerawala, S. M. & Hanafi, A. R. (1948). *Sankhyā*, **8** (1), 345.
MacMahon, P. A. (1882). *Amer. J. Math*, **5** (6), 289.
MacMahon, P. A. (1915). *Combinatory Analysis*, **1**. Cambridge University Press.
Sukhatme, P. V. (1938). *Philos. Trans.* A, **237**, 375.
Sukhatme, P. V. (1944). *Sankhyā*, **6**, 363.

Table 2.2

$w=2$	(1^2)	(2)
a_2	1	−2
a_1^2	2	1

Table 2.3

$w=3$	(1^3)	(21)	(3)
a_3	1	−3	3
a_2a_1	3	1	−3
a_1^3	6	3	1

Table 2.4

$w=4$	(1^4)	(21^2)	(2^2)	(31)	(4)
a_4	1	−4	2	4	−4
a_3a_1	4	1	−2	−1	4
a_2^2	6	2	1	−2	2
$a_2a_1^2$	12	5	2	1	−4
a_1^4	24	12	6	4	1

Table 2.5

$w=5$	(1^5)	(21^3)	(2^21)	(31^2)	(32)	(41)	(5)
a_5	1	−5	5	5	−5	−5	5
a_4a_1	5	1	−3	−1	5	1	−5
a_3a_2	10	3	1	−2	−1	5	−5
$a_3a_1^2$	20	7	2	1	−2	−1	5
$a_2^2a_1$	30	12	5	2	1	−3	5
$a_2a_1^3$	60	27	12	7	3	1	−5
a_1^5	120	60	30	20	10	5	1

Table 2.6

$w=6$	(1^6)	(21^4)	(2^21^2)	(31^3)	(2^3)	(321)	(41^2)	(3^2)	(42)	(51)	(6)
a_6	1	−6	9	6	−2	−12	−6	3	6	6	−6
a_5a_1	6	1	−4	−1	2	7	1	−3	−6	−1	6
a_4a_2	15	4	1	−2	−2	4	2	−3	2	−6	6
$a_4a_1^2$	30	9	2	1	.	−3	−1	3	2	1	−6
a_3^2	20	6	2	.	1	−3	3	3	−3	−3	3
$a_3a_2a_1$	60	22	8	3	3	1	−3	−3	4	7	−12
$a_3a_1^3$	120	48	18	10	6	3	1	.	−2	−1	6
a_2^3	90	36	15	6	6	3	.	1	−2	2	−2
$a_2^2a_1^2$	180	78	34	18	15	8	2	2	1	−4	9
$a_2a_1^4$	360	168	78	48	36	22	9	6	4	1	−6
a_1^6	720	360	180	120	90	60	30	20	15	6	1

Table 2.7

$w=7$	(1^7)	(21^5)	(2^21^3)	(31^4)	(2^31)	(321^2)	(41^3)	(32^2)	(3^21)	(421)	(51^2)	(43)	(52)	(61)	(7)
a_7	1	−7	14	7	−7	−21	−7	7	7	14	7	−7	−7	−7	7
a_6a_1	7	1	−5	−1	5	9	1	−7	−4	−8	−1	7	7	1	−7
a_5a_2	21	5	1	−2	−3	6	2	3	−7	−4	−2	7	−3	7	−7
$a_5a_1^2$	42	11	2	1	.	−4	−1	2	4	3	1	−7	−2	−1	7
a_4a_3	35	10	3	.	1	−3	3	−1	5	−2	−7	−5	7	7	−7
$a_4a_2a_1$	105	35	11	4	3	1	−3	−2	−1	8	3	−2	−4	−8	14
$a_4a_1^3$	210	75	24	13	6	3	1	.	.	−3	−1	3	2	1	−7
$a_3^2a_1$	140	50	18	6	7	2	.	1	−2	−1	4	5	−7	−4	7
$a_3a_2^2$	210	80	31	12	12	5	.	2	1	−2	2	−1	3	−7	7
$a_3a_2a_1^2$	420	170	68	34	27	13	3	5	2	1	−4	−3	6	9	−21
$a_3a_1^4$	840	360	150	88	60	34	13	12	6	4	1	.	−2	−1	7
$a_2^3a_1$	630	270	117	60	51	27	6	12	7	3	.	1	−3	5	−7
$a_2^2a_1^3$	1260	570	258	150	117	68	24	31	18	11	2	3	1	−5	14
$a_2a_1^5$	2520	1200	570	360	270	170	75	80	50	35	11	10	5	1	−7
a_1^7	5040	2520	1260	840	630	420	210	210	140	105	42	35	21	7	1

Table 2.8

$w=8$ (i)	(1^8)	(21^6)	(2^31^4)	(31^5)	(3^21^2)	(321^3)	(41^4)	(2^4)	(32^21)	(3^22^2)
a_8	1	−8	20	8	−16	−32	−8	2	24	12
a_7a_1	8	1	−6	−1	9	11	1	−2	−17	−5
a_6a_2	28	6	1	−2	−4	8	2	2	.	−9
$a_6a_1^2$	56	13	2	1	.	−5	−1	.	5	5
a_5a_3	56	15	4	.	1	−3	3	−2	6	3
$a_5a_2a_1$	168	51	14	5	1	1	−3	.	−3	−1
$a_5a_1^3$	336	108	30	10	6	3	1	.	.	.
a_4^2	70	20	6	.	2	.	.	1	−4	2
$a_4a_3a_1$	280	95	32	10	11	3	.	4	1	−2
$a_4a_2^2$	420	150	53	20	18	7	.	6	2	1
$a_4a_2a_1^2$	840	315	114	55	39	18	4	12	5	2
$a_4a_1^4$	1680	660	246	140	84	46	17	24	12	6
$a_3^2a_2$	560	210	80	30	31	12	.	12	5	2
$a_3^2a_1^2$	1120	440	172	80	68	30	6	28	12	4
$a_3a_2^2a_1$	1680	690	284	140	117	58	12	48	24	12
$a_3a_2a_1^3$	3360	1440	612	340	258	141	46	108	58	30
$a_3a_1^5$	6720	3000	1320	800	570	340	140	240	140	80
a_2^4	2520	1080	468	240	204	108	24	90	48	28
$a_2^3a_1^2$	5040	2250	1008	570	453	258	84	204	117	68
$a_2^2a_1^4$	10080	4680	2172	1320	1008	612	246	468	284	172
$a_2a_1^6$	20160	9720	4680	3000	2250	1440	660	1080	690	440
a_1^8	40320	20160	10080	6720	5040	3360	1680	2520	1680	1120

$w=8$ (ii)	(421^3)	(51^3)	(3^22)	(42^2)	(431)	(521)	(61^2)	(4^2)	(53)	(62)	(71)	(8)
a_8	24	8	−8	−8	−16	−16	−8	4	8	8	8	−8
a_7a_1	−10	−1	8	+8	9	9	1	−4	−8	−8	−1	8
a_6a_2	−6	−2	2	−4	16	4	2	−4	−8	4	−8	8
$a_6a_1^2$	4	1	−5	−2	−9	−3	−1	4	8	2	1	−8
a_5a_3	−9	−3	−7	8	1	1	8	−4	7	−8	−8	8
$a_5a_2a_1$	11	3	5	−4	−10	−8	−3	8	1	4	9	−16
$a_5a_1^3$	−4	−1	.	2	4	3	1	−4	−3	−2	−1	8
a_4^2	4	−4	4	−4	−8	8	4	6	−4	−4	−4	4
$a_4a_3a_1$	−1	4	−1	.	10	−10	−9	−8	1	10	9	−16
$a_4a_2^2$	−2	2	−2	4	.	−4	−2	−4	8	−4	8	−8
$a_4a_2a_1^2$	1	−4	.	−2	−1	11	4	4	−9	−6	−10	24
$a_4a_1^4$	4	1	.	.	.	−3	−1	.	3	2	1	−8
$a_3^2a_2$	.	.	1	−2	−1	5	−5	4	−7	2	8	−8
$a_3^2a_1^2$	2	.	2	1	−2	−1	5	2	3	−9	−5	12
$a_3a_2^2a_1$	5	.	5	2	1	−3	5	−4	6	.	−17	24
$a_3a_2a_1^3$	18	3	12	7	3	1	−5	.	−3	8	11	−32
$a_3a_1^5$	55	16	30	20	10	5	1	.	.	−2	−1	8
a_2^4	12	.	12	6	4	.	.	1	−2	2	−2	2
$a_2^3a_1^2$	39	6	31	18	11	3	.	2	1	−4	9	−16
$a_2^2a_1^4$	114	30	80	53	32	14	2	6	4	1	−6	20
$a_2a_1^6$	315	108	210	150	95	51	13	20	15	6	1	−8
a_1^8	840	336	560	420	280	168	56	70	56	28	8	1

Table 2.9

$w=9$ (i)	(1^9)	(21^7)	(2^21^5)	(31^6)	(2^21^3)	(321^4)	(41^5)	(2^41)	(32^21^2)	(3^21^3)	(421^3)	(51^4)	(32^3)
a_9	1	−9	27	9	−30	−45	−9	9	54	18	36	9	−9
a_8a_1	9	1	−7	−1	14	13	1	−7	−30	−6	−12	−1	9
a_7a_2	36	7	1	−2	−5	10	2	5	−5	−11	−8	−2	−5
$a_7a_1^2$	72	15	2	1	.	−6	−1	.	9	6	5	1	−2
a_6a_3	84	21	5	.	1	−3	3	−3	9	3	−12	−3	3
$a_6a_2a_1$	252	70	17	6	3	1	−3	.	−4	−1	14	3	2
$a_6a_1^3$	504	147	36	19	6	3	1	.	.	.	−5	−1	.
a_5a_4	126	35	10	.	3	.	.	1	−4	2	4	−4	−1
$a_5a_3a_1$	504	101	50	15	15	4	.	4	1	−2	−1	4	−2
$a_5a_2^2$	756	252	81	30	24	9	.	6	2	1	−2	2	.
$a_5a_2a_1^2$	1512	525	172	81	51	23	5	12	5	2	1	−4	.
$a_5a_1^4$	3024	1092	366	204	108	58	21	24	12	6	4	1	.
$a_4^2a_1$	630	210	70	20	24	6	.	9	2	.	.	.	1
$a_4a_3a_2$	1260	455	165	60	60	22	.	22	8	3	.	.	3
$a_4a_3a_1^2$	2520	945	350	155	129	54	10	48	19	6	3	.	7
$a_4a_2^2a_1$	3780	1470	565	270	213	101	20	78	36	17	7	.	12
$a_4a_2a_1^3$	7560	3045	1200	645	459	241	75	168	85	42	25	4	27
$a_4a_1^5$	15120	6300	2550	1500	990	570	225	360	200	110	75	21	60
a_3^3	1680	630	240	90	93	36	.	36	15	6	.	.	6
$a_3^2a_2a_1$	5040	2030	820	390	333	158	30	136	65	30	12	.	27
$a_3^2a_1^3$	10080	4200	1740	920	720	372	110	300	152	72	42	6	6
$a_3a_2^3$	7560	3150	1320	660	555	282	60	234	120	64	27	.	50
$a_3a_2^2a_1^2$	15120	6510	2800	1530	1203	656	200	516	281	152	85	12	12
$a_3a_2a_1^4$	30240	13440	5940	3480	2610	1516	570	1140	656	372	241	58	282
$a_3a_1^6$	60480	27720	12600	7800	5670	3480	1500	2520	1530	920	645	204	660
$a_2^4a_1$	22680	10080	4500	2520	2016	1140	360	906	516	300	168	24	234
$a_2^3a_1^3$	45360	20790	9540	5670	4383	2610	990	2016	1203	720	459	108	555
$a_2^2a_1^5$	90720	42840	20220	12600	9540	5940	2550	4500	2800	1740	1200	366	1320
$a_2a_1^7$	181440	88200	42840	27720	20790	13440	6300	10080	6510	4200	3045	1092	3150
a_1^9	362880	181440	90720	60480	45360	30240	15120	22680	15120	10080	7560	3024	7560

$w=9$ (ii)	(3^21)	(42^21)	(431^2)	(521^2)	(61^3)	(3^3)	(432)	(52^2)	(4^21)	(531)	(621)	(71^2)	(54)	(63)	(72)	(81)	(9)
a_9	−27	−27	−27	−27	−9	3	18	9	9	18	18	9	−9	−9	−9	−9	9
a_8a_1	19	19	11	11	1	−3	−18	−9	−5	−10	−10	−1	9	9	9	1	−9
a_7a_2	13	−1	20	6	2	−3	−4	5	−9	−18	−4	−2	9	9	−5	9	−9
$a_7a_1^2$	−12	−5	−11	−4	−1	3	11	2	5	10	3	1	−9	−9	−2	−1	9
a_6a_3	−18	9	9	9	3	6	.	−9	−9	.	.	−9	9	−9	9	9	−9
$a_6a_2a_1$	7	−12	−13	−11	−3	−3	4	4	14	10	8	3	−18	.	−4	−10	18
$a_6a_1^3$	.	5	5	4	1	.	−5	−2	−5	−4	−3	−1	9	3	2	1	−9
a_5a_4	7	−3	−13	7	9	−3	2	1	11	2	−18	−9	−11	9	9	9	−9
$a_5a_3a_1$	4	2	5	−15	−4	−3	−2	8	−6	5	10	10	2	.	−18	−10	18
$a_5a_2^2$	−3	6	.	−6	−2	3	−6	.	−1	8	4	2	1	−9	5	−9	9
$a_5a_2a_1^2$	.	−3	−1	15	4	.	5	−6	1	−15	−11	−4	7	9	6	11	−27
$a_5a_1^4$	.	.	.	−4	−1	.	.	2	.	4	3	1	−4	−3	−2	−1	9
$a_4^2a_1$	−3	−1	5	1	−5	3	−2	−1	−5	−6	14	5	11	−9	−9	−5	9
$a_4a_3a_2$	1	−2	−1	5	−5	−3	8	−6	−2	−2	4	11	2	.	−4	−18	18
$a_4a_3a_1^2$	2	1	−2	−1	5	.	−1	.	5	5	−13	−11	−13	9	20	11	−27
$a_4a_2^2a_1$	5	2	1	−3	5	.	−2	6	−1	2	−12	−5	−3	9	−1	19	−27
$a_4a_2a_1^3$	12	7	3	1	−5	.	.	−2	.	−1	14	5	4	−12	−8	−12	36
$a_4a_1^5$	30	20	10	5	1	.	.	.	.	.	−3	−1	.	3	2	1	−9
a_3^3	3	.	.	.	.	1	−3	3	3	−3	−3	3	−3	6	−3	−3	3
$a_3^2a_2a_1$	13	5	2	.	.	3	1	−3	−3	4	7	−12	7	−18	13	19	−27
$a_3^2a_1^3$	30	17	6	2	.	6	3	1	.	−2	−1	6	2	3	−1	−6	18
$a_3a_2^3$	27	12	7	.	.	6	3	.	1	−2	2	−2	−1	3	−5	9	−9
$a_3a_2^2a_1^2$	65	36	19	5	.	15	8	2	2	1	−4	9	−4	9	−5	−30	54
$a_3a_2a_1^4$	158	101	54	23	3	36	22	9	6	4	1	−6	.	−3	10	13	−45
$a_3a_1^6$	390	270	155	81	19	90	60	30	20	15	6	1	.	.	−2	−1	9
$a_2^4a_1$	136	78	48	12	.	36	22	6	9	4	.	.	1	−3	5	−7	9
$a_2^3a_1^3$	333	213	129	51	6	93	60	24	24	15	3	.	3	1	−5	14	−30
$a_2^2a_1^5$	820	565	350	172	36	240	165	81	70	50	17	2	10	5	1	−7	+27
$a_2a_1^7$	2030	1470	945	525	147	630	455	252	210	161	70	15	35	21	7	1	−9
a_1^9	5040	3780	2520	1512	504	1680	1260	756	630	504	252	72	126	84	36	9	1

Table 2.10

$w=10$ (i)	(1^{10})	(21^8)	(2^21^6)	(31^7)	(2^21^4)	(321^5)	(41^6)	(2^41^2)	(32^21^3)
a_{10}	1	−10	35	10	−50	−60	−10	25	100
a_9a_1	10	1	−8	−1	20	15	1	−16	−46
a_8a_2	45	8	1	−2	−6	12	2	9	−12
$a_8a_1^2$	90	17	2	1	.	−7	−1	.	14
a_7a_3	120	28	6	.	1	−3	3	−4	12
$a_7a_2a_1$	360	92	20	7	3	1	−3	.	−5
$a_7a_1^3$	720	192	42	22	6	3	1	.	.
a_6a_4	210	56	15	.	4	.	.	1	−4
$a_6a_3a_1$	840	252	72	21	19	5	.	4	1
$a_6a_2^2$	1260	392	115	42	30	11	.	6	2
$a_6a_2a_1^2$	2520	812	242	112	63	28	6	12	5
$a_6a_1^4$	5040	1680	510	280	132	70	25	24	12
a_5^2	252	70	20	.	6	.	.	2	.
$a_5a_4a_1$	1200	406	130	35	42	10	.	14	3
$a_5a_3a_2$	2520	868	296	105	99	35	.	32	11
$a_5a_3a_1^2$	5040	1792	622	266	210	85	15	68	26
$a_5a_2^2a_1$	7560	2772	990	462	339	156	30	108	48
$a_5a_2a_1^3$	15120	5712	2082	1092	720	368	111	228	112
$a_5a_1^5$	30240	11760	4380	2520	1530	860	330	480	260
$a_4^2a_2$	3150	1120	400	140	144	50	.	53	18
$a_4^2a_1^2$	6300	2310	840	350	306	120	20	114	42
$a_4a_3^2$	4200	1540	570	210	213	80	.	80	31
$a_4a_3a_2a_1$	12600	4900	1900	875	735	335	60	284	128
$a_4a_3a_1^3$	25200	10080	3990	2030	1566	775	215	612	294
$a_4a_2^3$	18900	7560	3015	1470	1194	585	120	468	228
$a_4a_2^2a_1^2$	37800	15540	6330	3360	2547	1340	390	1008	523
$a_4a_2a_1^4$	75600	31920	13290	7560	5436	3050	1095	2172	1196
$a_4a_1^6$	151200	65520	27900	16800	11610	6900	2850	4680	2730
$a_3^3a_1$	16800	6720	2700	1260	1092	510	90	444	210
$a_3^2a_2^2$	25200	10360	4280	2100	1776	880	180	740	370
$a_3^2a_2a_1^2$	50400	21280	8980	4760	3792	2000	570	1604	844
$a_3^2a_1^4$	100800	43680	18840	10640	8100	4520	1580	3480	1920
$a_3a_2^3a_1$	75600	32760	14220	7770	6180	3390	1020	2688	1479
$a_3a_2^2a_1^3$	151200	67200	29820	17220	13212	7610	2730	5844	3358
$a_3a_2a_1^5$	302400	137760	62520	37800	28260	17000	6900	12720	7610
$a_3a_1^7$	604800	282240	131040	82320	60480	37800	16800	27720	17220
a_2^5	113400	50400	22500	12600	10080	5700	1800	4530	2580
$a_2^4a_1^2$	226800	103320	47160	27720	21564	12720	4680	9876	5844
$a_2^3a_1^4$	453600	211680	98820	60480	46152	28260	11610	21564	13212
$a_2^2a_1^6$	907200	433440	207000	131040	98820	62520	27900	47160	29820
$a_2a_1^8$	1814400	887040	433440	282240	211680	137760	65520	103320	67200
a_1^{10}	3628800	1814400	907200	604800	453600	302400	151200	226800	151200

$w=10$ (ii)	(3^21^4)	(421^4)	(51^5)	(2^5)	(32^21)	(3^221^2)	(42^21^2)	(431^3)	(521^3)	(61^4)
a_{10}	25	50	10	−2	−40	−60	−60	−40	−40	−10
a_9a_1	−7	−14	−1	2	31	33	33	13	13	1
a_8a_2	−13	−10	−2	−2	−8	28	4	24	8	2
$a_8a_1^2$	7	6	1	.	−7	−21	−9	−13	−5	−1
a_7a_3	3	−15	−3	2	−2	−24	18	12	12	3
$a_7a_2a_1$	−1	17	3	.	5	9	−23	−16	−14	−3
$a_7a_1^3$	.	−6	−1	.	.	.	9	6	5	1
a_6a_4	2	4	−4	−2	8	.	−12	−8	16	4
$a_6a_3a_1$	−2	−1	4	.	−3	6	3	5	−10	−4
$a_6a_2^2$	1	−2	2	.	.	−4	8	.	−8	−2
$a_6a_2a_1^2$	2	1	−4	.	.	.	−4	−1	10	4
$a_6a_1^4$	6	4	1	.	.	.	.	.	−5	−1
a_5^2	.	.	.	1	−5	5	5	−5	−5	5
$a_5a_4a_1$	.	.	.	5	1	−3	−1	5	1	−5
$a_5a_3a_2$	4	.	.	10	3	1	−2	−1	5	−5
$a_5a_3a_1^2$	8	4	.	20	7	2	1	−2	−1	+5
$a_5a_2^2a_1$	22	9	.	30	12	5	2	1	−3	5
$a_5a_2a_1^3$	54	32	5	60	27	12	7	3	1	−5
$a_5a_1^5$	140	95	26	120	60	30	20	10	5	1
$a_4^2a_2$	6	.	.	20	7	2	.	.	.	.
$a_4^2a_1^2$	12	6	.	45	16	4	2	.	.	.
$a_4a_3^2$	12	.	.	30	12	5	.	.	.	.
$a_4a_3a_2a_1$	56	22	.	110	49	21	8	3	.	.
$a_4a_3a_1^3$	132	76	10	240	112	48	27	9	3	.
$a_4a_2^3$	115	48	.	180	87	42	18	10	.	.
$a_4a_2^2a_1^2$	270	140	20	390	198	99	54	27	7	.
$a_4a_2a_1^4$	650	416	95	840	450	236	140	76	32	4
$a_4a_1^6$	1580	1095	330	1800	1020	570	390	215	111	25
$a_3^3a_2$	96	36	.	180	87	42	15	6	.	.
$a_3^3a_2^2$	188	78	.	310	156	80	34	18	.	.
$a_3^3a_2a_1^2$	436	236	30	680	358	186	99	48	12	.
$a_3^3a_1^4$	1032	650	140	1500	820	436	270	132	54	6
$a_3a_2^4a_1$	820	450	60	1170	645	358	198	112	27	.
$a_3a_2^3a_1^3$	1920	1196	260	2580	1479	844	523	294	112	12
$a_3a_2a_1^5$	4520	3050	860	5700	3390	2000	1340	775	368	70
$a_3a_1^7$	10640	7560	2520	12600	7770	4760	3360	2030	1092	280
a_2^6	1500	840	120	2040	1170	680	390	240	60	.
$a_2^5a_1^2$	3480	2172	480	4530	2688	1604	1008	612	228	24
$a_2^4a_1^4$	8100	5436	1530	10080	6180	3792	2547	1566	720	132
$a_2^3a_1^6$	18840	13290	4380	22500	14220	8980	6330	3990	2082	510
$a_2a_1^8$	43680	31920	11760	50400	32760	21280	15540	10080	5712	1680
a_1^{10}	100800	75600	30240	113400	75600	50400	37800	25200	15120	5040

Table 2.10 (*cont.*)

$w = 10$ (iii)	$(3^2 2^2)$	(42^2)	$(3^2 1)$	(4321)	$(52^2 1)$	$(4^2 1^2)$	(531^2)	(621^2)	(71^3)	(43^2)	$(4^2 2)$
a_{10}	15	10	10	60	30	15	30	30	10	−10	−10
$a_9 a_1$	−15	−10	−7	−42	−21	−6	−12	−12	−1	10	10
$a_8 a_2$	1	6	−10	−28	2	11	−22	−6	−2	10	2
$a_8 a_1^2$	7	2	7	26	5	6	12	4	1	−10	−6
$a_7 a_3$	6	−10	11	3	−9	−15	−9	−9	−3	−11	10
$a_7 a_2 a_1$	−7	4	−4	21	12	17	13	11	3	1	−12
$a_7 a_1^3$	.	−2	.	−12	−5	−6	−5	−4	−1	3	6
$a_6 a_4$	−9	10	2	12	−18	9	−6	−6	−10	−2	14
$a_6 a_3 a_1$	3	.	−7	−15	18	−3	15	15	4	13	4
$a_6 a_2^2$	2	−4	4	−8	4	−1	10	6	2	−4	10
$a_6 a_1^4$	.	.	.	.	.	.	.	.	.	.	.
$a_5 a_3 a_1^2$	.	2	.	7	−17	1	−19	−15	−4	−3	−6
$a_5 a_2^2 a_1$	.	.	.	.	5	.	5	4	1	.	.
a_5^2	5	−5	−5	−5	10	5	10	−15	−5	5	5
$a_5 a_4 a_1$	−1	.	5	.	−1	−8	−12	18	11	−8	4
$a_5 a_3 a_2$	−2	4	−1	5	−13	1	−4	15	5	1	−12
$a_5 a_2 a_1^3$	.	−2	.	4	2	2	3	−19	−5	−3	2
$a_5 a_2^3 a_1$	.	.	.	−3	9	−1	2	−17	−5	3	2
$a_5 a_2 a_1^3$	.	.	.	.	−3	.	−1	19	5	.	.
$a_5 a_1^5$	.	.	.	.	.	.	.	−4	−1	.	.
$a_4^2 a_2$	1	−2	−2	4	2	−3	2	−6	6	2	2
$a_4^2 a_1^2$	2	1	.	−3	−1	3	2	1	−6	3	−3
$a_4 a_2^3$	2	.	1	−3	3	3	−3	−3	3	−1	2
$a_4 a_3 a_2 a_1$	8	3	3	1	−3	−3	4	7	−12	−3	4
$a_4 a_3 a_1^3$	18	10	6	3	1	.	−2	−1	6	.	.
$a_4 a_2^3 a_1$	15	6	6	8	.	1	−2	2	−2	.	−2
$a_4 a_2^2 a_1^2$	34	18	15	8	2	2	1	−4	9	.	.
$a_4 a_2 a_1^4$	78	48	30	22	9	6	4	1	−6	.	.
$a_4 a_1^6$	180	120	90	60	30	20	15	6	1	.	.
$a_3^3 a_1$	18	6	10	3	.	.	.	.	.	1	−2
$a_3^2 a_2^2$	34	15	18	8	.	2	.	.	.	2	1
$a_3^2 a_2 a_1^2$	80	42	42	21	5	4	2	.	.	5	2
$a_3^2 a_1^4$	188	115	96	50	22	12	8	2	.	12	6
$a_3 a_2^3 a_1$	156	87	87	49	12	16	7	.	.	12	7
$a_3 a_2^2 a_1^3$	370	228	210	128	48	42	26	5	.	31	18
$a_3 a_2 a_1^5$	880	585	510	335	156	120	85	28	3	80	50
$a_3 a_1^7$	2100	1470	1260	875	462	350	266	112	22	210	140
a_2^5	310	180	180	110	30	45	20	.	.	30	20
$a_2^4 a_1^2$	740	468	444	284	108	114	68	12	.	80	53
$a_2^3 a_1^4$	1776	1194	1092	735	339	306	210	63	6	213	144
$a_2^2 a_1^6$	4280	3015	2700	1900	990	840	622	242	42	570	400
$a_2 a_1^8$	10360	7560	6720	4900	2772	2310	1792	812	192	1540	1120
a_1^{10}	25200	18900	16800	12600	7560	6300	5040	2520	720	4200	3150

$w = 10$ (iv)	(532)	(62^2)	(541)	(631)	(721)	(81^2)	(5^2)	(64)	(73)	(82)	(91)	(10)
a_{10}	−20	−10	−20	−20	−20	−10	5	10	10	10	10	−10
$a_9 a_1$	20	10	11	11	11	1	−5	−10	−10	−10	−1	10
$a_8 a_2$	4	−6	20	20	4	2	−5	−10	−10	6	−10	10
$a_8 a_1^2$	−12	−2	−11	−11	−3	−1	5	10	10	2	1	−10
$a_7 a_3$	−1	10	20	−1	−1	10	−5	−10	11	−10	−10	10
$a_7 a_2 a_1$	−3	−4	−31	−10	−8	−3	10	20	−1	4	11	−20
$a_7 a_1^3$	5	2	11	4	3	1	−5	−10	−3	−2	−1	10
$a_6 a_4$	20	−2	−4	−4	20	10	−5	14	−10	−10	−10	10
$a_6 a_3 a_1$	−19	−8	−7	−4	−10	−11	10	−4	−1	20	11	−20
$a_6 a_2^2$	−4	.	−8	−8	−4	−2	5	−2	10	−6	10	−10
$a_6 a_2 a_1^3$	15	6	18	15	11	4	−15	−6	−9	−6	−12	30
$a_6 a_1^4$	−5	−2	−5	−4	−3	−1	5	4	3	2	1	−10
a_5^2	−15	5	−15	10	10	5	10	−5	−5	−5	−5	5
$a_5 a_4 a_1$	10	−8	23	−7	−31	−11	−15	−4	20	20	11	−20
$a_5 a_3 a_2$	17	−4	10	−19	−3	−12	−15	20	−1	4	20	−20
$a_5 a_3 a_1^2$	−4	10	−12	15	13	12	10	−6	−9	−22	−12	30
$a_5 a_2^2 a_1$	−13	4	−1	18	12	5	10	−18	−9	2	−21	30
$a_5 a_2 a_1^3$	5	−8	1	−19	−14	−5	−5	16	12	8	13	−40
$a_5 a_1^5$	.	2	.	4	3	1	.	−4	−3	−2	−1	10
$a_4^2 a_2$	−12	10	4	4	−12	−6	5	−14	10	2	10	−10
$a_4^2 a_1^3$	1	−1	−8	−3	17	6	5	9	−15	−11	−6	15
$a_4 a_3^2$	1	−4	−8	13	1	−10	5	−2	−11	10	10	−10
$a_4 a_3 a_2 a_1$	5	−8	.	−15	21	26	−5	12	3	−28	−42	60
$a_4 a_3 a_1^3$	−1	.	5	5	−16	−13	−5	−8	12	24	13	−40
$a_4 a_2^3$	4	−4	.	.	4	2	−5	10	−10	6	−10	10
$a_4 a_2^2 a_1^2$	−2	8	−1	3	−23	−9	5	−12	18	4	33	−60
$a_4 a_2 a_1^4$	.	−2	.	−1	17	6	.	4	−15	−10	−14	50
$a_4 a_1^6$	.	.	.	.	−3	−1	.	.	3	2	1	−10
$a_3^3 a_2$	−1	4	5	−7	−4	7	−5	2	11	−10	−7	10
$a_3^2 a_2^3$	−2	2	−1	3	−7	7	5	−9	6	1	−15	15
$a_3^2 a_2 a_1^2$	1	−4	−3	6	9	−21	5	.	−24	28	33	−60
$a_3^2 a_1^4$	4	1	.	−2	−1	7	.	2	3	−13	−7	25
$a_3 a_2^3 a_1$	3	.	1	−3	5	−7	−5	8	−2	−8	31	−40
$a_3 a_2^2 a_1^3$	11	2	3	1	−5	14	.	−4	12	−12	−46	100
$a_3 a_2 a_1^5$	35	11	10	5	1	−7	.	.	−3	12	15	−60
$a_3 a_1^7$	105	42	35	21	7	1	.	.	.	−2	−1	10
a_2^5	10	.	5	.	.	.	1	−2	2	−2	2	−2
$a_2^4 a_1^2$	32	6	14	4	.	.	2	1	−4	9	−16	25
$a_2^3 a_1^4$	99	30	42	19	3	.	6	4	1	−6	20	−50
$a_2^2 a_1^6$	296	115	130	72	20	2	20	15	6	1	−8	35
$a_2 a_1^8$	868	392	406	252	92	17	70	56	28	8	1	−10
a_1^{10}	2520	1260	1260	840	360	90	252	210	120	45	10	1

Table 2.11

$w = 11$ (i)	(1^{11})	(21^9)	(2^21^7)	(31^8)	(2^31^5)	(321^4)	(41^7)	(2^41^3)	(32^31^4)	(3^21^5)
a_{11}	1	−11	44	11	−77	−77	−11	55	165	33
$a_{10}a_1$	11	1	−9	−1	27	17	1	−30	−65	−8
a_9a_2	55	9	1	−2	−7	14	2	14	−21	−15
$a_9a_1^2$	110	19	2	1	.	−8	−1	.	20	8
a_8a_3	165	36	7	.	1	−3	3	−5	15	3
$a_8a_2a_1$	495	117	23	8	3	1	−3	.	−6	−1
$a_8a_1^3$	990	243	48	25	6	3	1	.	.	.
a_7a_4	330	84	21	.	5	.	.	1	−4	2
$a_7a_3a_1$	1320	372	98	28	23	6	.	4	1	−2
$a_7a_2^2$	1980	576	155	56	36	13	.	6	2	1
$a_7a_2a_1^2$	3960	1188	324	148	75	33	7	12	5	2
$a_7a_1^4$	7920	2448	678	368	156	82	29	24	12	6
a_6a_5	462	126	35	.	10	.	.	3	.	.
$a_6a_4a_1$	2310	714	217	56	65	15	.	19	4	.
$a_6a_3a_2$	4620	1512	483	168	148	51	.	42	14	5
$a_6a_3a_1^2$	9240	3108	1008	420	311	123	21	88	33	10
$a_6a_2^2a_1$	13860	4788	1589	728	495	223	42	138	60	27
$a_6a_2a_1^3$	27720	9828	3318	1708	1041	522	154	288	139	66
$a_6a_1^5$	55440	20160	6930	3920	2190	1210	455	600	320	170
$a_5^2a_1$	2772	882	280	70	90	20	.	30	6	.
$a_5a_4a_2$	6930	2394	826	280	285	95	.	99	32	10
$a_5a_4a_1^2$	13860	4914	1722	686	600	225	35	210	74	20
$a_5a_3^2$	9240	3276	1162	420	411	150	.	144	53	20
$a_5a_3a_2a_1$	27720	10332	3808	1708	1383	611	105	492	213	90
$a_5a_3a_1^3$	55440	21168	7938	3920	2916	1398	371	1044	484	210
$a_5a_2^3$	41580	15876	5985	2856	2211	1053	210	792	372	181
$a_5a_2^2a_1^2$	83160	32508	12474	6468	4665	2388	672	1680	843	422
$a_5a_2a_1^4$	166320	66528	25998	14448	9846	5382	1869	3564	1904	1006
$a_5a_1^6$	332640	136080	54180	31920	20790	12060	4830	7560	4290	2420
$a_4^2a_3$	11550	4200	1540	560	570	210	.	213	80	30
$a_4^2a_3a_1$	34650	13230	5040	2240	1920	840	140	735	316	130
$a_4^2a_2^2$	69300	27090	10500	5110	4050	1910	490	1566	714	300
$a_4^2a_2a_1^2$	46200	18060	7070	3220	2775	1260	210	1092	497	220
$a_4a_3^2a_2$	69300	27720	11095	5320	4440	2135	420	1776	854	415
$a_4a_3^2a_1^3$	138600	56700	23100	11900	9375	4785	1295	3792	1917	930
$a_4a_3a_2^2$	277200	115920	48090	26320	19800	10670	3535	8100	4292	2210
$a_4a_3a_2a_1^3$	207900	86940	36225	19320	15015	7995	2310	6180	3276	1745
$a_4a_2^3a_1^2$	415800	177660	75390	42420	31725	17730	6090	13212	7319	4030
$a_4a_2a_1^5$	831600	362880	156870	92400	67050	39150	15225	28260	16320	9350
$a_4a_1^7$	1663200	740880	326340	199920	141750	86100	36750	60480	36330	21700
$a_3^3a_2$	92400	37800	15540	7560	6420	3150	630	2664	1320	660
$a_3^3a_1^2$	184800	77280	32340	16800	13560	7020	1890	5700	2952	1500
$a_3^2a_2^2a_1$	277200	118440	50680	27160	21720	11660	3360	9324	5016	2700
$a_3^2a_2a_1^3$	554400	241920	105420	59360	45900	25740	8750	19980	11168	6180
$a_3^2a_1^5$	1108800	493920	219240	128800	97020	56600	21700	42840	24820	14200
$a_3a_2^4$	415800	181440	79380	43680	34800	19260	5880	15282	8496	4780
$a_3a_2^3a_1^2$	831600	370440	165060	94920	73560	42330	14910	32784	18876	10880
$a_3a_2^2a_1^4$	1663200	756000	343140	204960	155520	92700	36330	70380	41864	24820
$a_3a_2a_1^6$	3326400	1542240	713160	440160	328860	202320	86100	151200	92700	56600
$a_3a_1^8$	6652800	3144960	1481760	940800	695520	440160	199920	325080	204960	128800
$a_2^5a_1$	1247400	567000	258300	151200	117000	69300	25200	53910	31800	18900
$a_2^4a_1^3$	2494800	1156680	536760	325080	249300	151200	60480	115884	70380	42840
$a_2^3a_1^5$	4989600	2358720	1115100	695520	527220	328860	141750	249300	155520	97020
$a_2^2a_1^7$	9979200	4808160	2315880	1481760	1115100	713160	326340	536760	343140	219240
$a_2a_1^9$	19958400	9797760	4808160	3144960	2358720	1542240	740880	1156680	756000	493920
a_1^{11}	39916800	19958400	9979200	6652800	4989600	3326400	1663200	2494800	1663200	1108800

Table 2.11 (*cont.*)

$w=11$ (ii)	(421^6)	(51^6)	(2^21)	(32^31^2)	(3^221^3)	(42^21^3)	(431^6)	(521^6)	(61^5)	(32^4)
a_{11}	66	11	−11	−110	−110	−110	−55	−55	−11	11
$a_{10}a_1$	−16	−1	9	70	50	50	15	15	1	−11
a_9a_2	−12	−2	−7	−7	47	11	28	10	2	7
$a_9a_1^2$	7	1	.	−16	−32	−14	−15	−6	−1	2
a_8a_3	−18	−3	5	−10	−30	30	15	15	3	−5
$a_8a_2a_1$	20	3	.	9	11	−37	−19	−17	−3	−2
$a_8a_1^3$	−7	−1	.	.	.	14	7	6	1	.
a_7a_4	4	−4	−3	12	−2	−16	−8	20	4	3
$a_7a_3a_1$	−1	4	.	−4	8	4	5	−23	−4	2
$a_7a_2^2$	−2	2	.	.	−5	10	.	−10	−2	.
$a_7a_2a_1^2$	1	−4	.	.	.	−5	−1	23	4	.
$a_7a_1^4$	4	1	.	.	.	.	.	−6	−1	.
a_6a_5	.	.	1	−5	5	5	−5	−5	5	−1
$a_6a_4a_1$	.	.	5	1	−3	−1	5	1	−5	−2
$a_6a_3a_2$	.	.	10	3	1	−2	−1	5	−5	.
$a_6a_3a_1^2$	5	.	20	7	2	1	−2	−1	5	.
$a_6a_2^2a_1$	11	.	30	12	5	2	1	−3	5	.
$a_6a_2a_1^3$	39	6	60	27	12	7	3	1	−5	.
$a_6a_1^5$	115	31	120	60	30	20	10	5	1	.
$a_5^2a_1$	.	.	11	2	.	.	.	.	.	1
$a_5a_4a_3$	.	.	35	11	3	.	.	.	.	4
$a_5a_4a_1^2$	10	.	75	25	6	3	.	.	.	9
$a_5a_3^2$	.	.	50	18	7	.	.	.	.	6
$a_5a_3a_2a_1$	35	.	170	71	29	11	4	.	.	22
$a_5a_3a_1^3$	120	15	360	160	66	37	12	4	.	48
$a_5a_2^3$	75	.	270	123	57	24	13	.	.	36
$a_5a_2^2a_1^2$	231	30	570	276	133	72	35	9	.	78
$a_5a_2a_1^4$	639	141	1200	618	314	197	98	41	5	168
$a_5a_1^6$	1665	486	2520	1380	750	510	275	141	31	360
$a_4^2a_3$	.	.	80	31	12	.	.	.	.	12
$a_4^2a_2a_1$	50	.	285	121	48	18	6	.	.	48
$a_4^2a_1^3$	170	20	615	272	108	60	18	6	.	109
$a_4a_3^2a_1$	80	.	430	197	89	31	12	.	.	78
$a_4a_3a_2^2$	170	.	710	341	164	68	34	.	.	136
$a_4a_3a_2a_1^3$	505	60	1530	766	375	196	90	22	.	306
$a_4a_2^3a_1$	1370	275	3300	1718	864	527	244	98	10	688
$a_4a_2^2a_1^3$	945	120	2520	1326	697	378	203	48	.	528
$a_4a_2a_1^5$	2475	510	5430	2973	1613	983	527	197	20	1182
$a_4a_1^7$	6225	1665	11700	6660	3750	2475	1370	639	115	2640
$a_3^3a_2$	15225	4830	25200	14910	8750	6090	3535	1869	455	5880
$a_3^3a_1$	270	.	1110	555	282	117	60	.	.	234
$a_3^2a_2^2$	780	90	2400	1248	642	327	156	36	.	528
$a_3^2a_2^2a_1$	1450	180	4010	2162	1168	628	338	78	.	934
$a_3^2a_2a_1^3$	3750	750	8700	4852	2682	1613	864	314	30	2112
$a_3^2a_1^5$	9350	2420	18900	10880	6180	4030	2210	1006	170	4780
$a_3a_2^4$	2040	360	6720	3750	2112	1182	688	168	.	1656
$a_3a_2^3a_1^2$	6660	1380	14610	8415	4852	2973	1718	618	60	3750
$a_3a_2^2a_1^4$	16320	4290	31800	18870	11168	7319	4292	1904	320	8496
$a_3a_2a_1^6$	39150	12060	69300	42330	25740	17730	10670	5382	1210	19260
$a_3a_1^8$	92400	31920	151200	94920	59360	42420	26320	14448	3920	43680
$a_2^5a_1$	11700	2520	24690	14610	8700	5430	3300	1200	120	6720
$a_2^4a_1^3$	28260	7560	53910	32784	19980	13212	8100	3564	600	15282
$a_2^3a_1^5$	67050	20790	117900	73560	45900	31725	19800	9840	2190	34800
$a_2^2a_1^7$	156870	54180	258300	165060	105420	75390	48090	25998	6930	79380
$a_2a_1^9$	362880	136080	567000	370440	241920	177660	115920	66528	20160	181440
a_1^{11}	831600	332640	1247400	831600	554400	415800	277200	166320	55440	415800

Table 2.11 (*cont.*)

$w=11$ (iii)	$(3^3 2^2 1)$	$(42^3 1)$	$(3^2 1^5)$	(4321^5)	$(52^2 1^5)$	$(4^2 1^3)$	(531^5)	(621^5)	(71^4)	$(3^2 2)$	(432^2)	(52^2)
a_{11}	66	44	22	132	66	22	44	44	11	−11	−33	−11
$a_{10}a_1$	−51	−34	−12	−72	−36	−7	−14	−14	−1	11	33	11
a_9a_2	−12	10	−19	−60	−3	−13	−20	−8	−2	5	−3	−7
$a_9a_1^2$	24	7	12	45	9	7	14	5	1	−8	−15	−2
a_8a_3	30	−20	14	−12	−18	−18	−12	−12	−3	−13	9	11
$a_8a_2a_1$	−17	16	−5	44	23	20	16	14	3	8	−6	−4
$a_8a_1^3$	.	−7	.	−21	−9	−7	−6	−5	−1	.	7	2
a_7a_4	−24	12	6	36	−24	6	−16	−10	−4	11	−9	−3
$a_7a_3a_1$	.	−2	−9	−21	33	−3	19	19	4	2	.	−8
$a_7a_2^2$	5	−10	5	−10	10	−1	12	8	2	−5	10	.
$a_7a_2a_1^2$	.	5	.	9	−32	1	−23	−19	−4	.	−7	6
$a_7a_1^4$	.	.	.	.	9	.	6	5	1	.	.	−2
a_6a_5	9	−4	−7	−12	9	8	16	−14	−11	−4	3	1
$a_6a_4a_1$	6	2	3	−12	−3	−5	−6	24	5	−7	6	2
$a_6a_3a_2$	−3	6	−1	6	−18	1	−4	20	5	5	−12	6
$a_6a_3a_1^2$	.	−3	.	6	3	2	3	−24	−5	.	3	.
$a_6a_2^2a_1$	.	.	.	−4	12	−1	2	−22	−5	.	2	−6
$a_6a_2a_1^3$	.	.	.	.	−4	.	−1	24	5	.	.	2
$a_6a_1^5$	.	.	.	.	.	.	.	−5	−1	.	.	.
$a_5^2a_1$	−4	−1	2	7	1	−3	−6	−1	6	4	−3	−1
$a_5a_4a_2$	1	−2	−2	4	2	−3	2	−6	6	−1	2	.
$a_5a_4a_1^2$	2	1	.	−3	−1	3	2	1	−6	.	−1	.
$a_5a_3^2$	2	.	1	−3	3	3	−3	−3	3	−2	6	−6
$a_5a_3a_2a_1$	8	3	3	1	−3	−3	4	7	−12	.	−2	6
$a_5a_3a_1^3$	18	10	6	3	1	.	−2	−1	6	.	.	−2
$a_5a_2^3$	15	6	6	3	.	1	−2	2	−2	.	.	.
$a_5a_2^2a_1^2$	34	18	15	8	2	2	1	−4	9	.	.	.
$a_5a_2a_1^4$	78	48	36	22	9	6	4	1	−6	.	.	.
$a_5a_1^6$	180	120	90	60	30	20	15	6	1	.	.	.
$a_4^2a_3$	5	.	2	.	.	.	.	.	.	1	−3	3
$a_4^2a_2a_1$	19	7	6	2	.	.	.	.	.	3	1	−3
$a_4^2a_1^3$	42	23	12	6	2	.	.	.	.	6	3	1
$a_4a_3^2a_1$	36	12	17	5	.	.	.	.	.	7	2	.
$a_4a_3a_2^2$	65	27	30	13	.	3	.	.	.	12	5	.
$a_4a_3a_2a_1^3$	148	76	69	34	8	6	3	.	.	27	13	3
$a_4a_3a_1^4$	338	203	156	90	35	18	12	3	.	60	34	13
$a_4a_2^3a_1$	273	147	138	76	18	23	10	.	.	51	27	6
$a_4a_2^2a_1^3$	628	378	327	196	72	60	37	7	.	117	68	24
$a_4a_2a_1^5$	1450	945	780	505	231	170	120	39	4	270	170	75
$a_4a_1^7$	3360	2310	1890	1295	672	490	371	154	29	630	420	210
$a_3^3a_2$	120	51	64	27	.	6	.	.	.	27	12	.
$a_3^3a_1^2$	276	138	146	69	15	12	6	.	.	64	30	6
$a_3^2a_2^2a_1$	506	273	276	148	34	42	18	.	.	120	65	15
$a_3^2a_2a_1^3$	1168	697	642	375	133	108	66	12	.	282	164	57
$a_3^2a_1^5$	2700	1745	1500	950	422	300	210	66	6	660	415	181
$a_3a_2^3$	934	528	528	306	78	109	48	.	.	234	136	36
$a_3a_2^3a_1^2$	2162	1326	1248	766	276	272	160	27	.	555	341	123
$a_3a_2^2a_1^4$	5016	3276	2952	1917	843	714	484	139	12	1320	854	372
$a_3a_2a_1^6$	11660	7995	7020	4785	2388	1910	1398	522	82	3150	2135	1053
$a_3a_1^8$	27160	19320	16800	11900	6468	5110	3920	1708	368	7560	5320	2856
$a_2^5a_1$	4010	2520	2400	1530	570	615	360	60	.	1110	710	270
$a_2^4a_1^3$	9324	6180	5700	3792	1680	1566	1044	288	24	2664	1776	792
$a_2^3a_1^5$	21720	15015	13560	9375	4665	4050	2916	1041	156	6420	4440	2211
$a_2^2a_1^7$	50680	36225	32340	23100	12474	10500	7938	3318	678	15540	11095	5985
$a_2a_1^9$	118440	86940	77280	56700	32508	27090	21168	9828	2448	37800	27720	15876
a_1^{11}	277200	207900	184800	138600	83160	69300	55440	27720	7920	92400	69300	41580

Table 2.11 (*cont.*)

$w = 11$ (iv)	$(43^2 1)$	$(4^2 21)$	(5321)	$(62^2 1)$	(541^2)	(631^2)	(721^2)	(81^3)	$(4^2 3)$	(53^2)	(542)	(632)
a_{11}	−33	−33	−66	−33	−33	−33	−33	−11	11	11	22	22
$a_{10}a_1$	23	23	46	23	13	13	13	1	−11	−11	−22	−22
a_9a_2	33	15	30	−3	24	24	6	2	−11	−11	−4	−4
$a_9a_1^2$	−23	−14	−28	−5	−13	−13	−4	−1	11	11	13	13
a_8a_3	−15	21	−6	9	33	9	9	3	1	13	−22	2
$a_8a_2a_1$	−8	−34	−20	−12	−37	−13	−11	−3	10	−2	26	2
$a_8a_1^3$	7	14	12	5	13	5	4	1	−7	−3	−13	−5
a_7a_4	−23	−23	38	10	5	5	5	11	17	−11	6	−22
$a_7a_3a_1$	27	12	−33	−18	−18	−15	−15	−4	−18	−2	16	20
$a_7a_2^2$	−5	13	−16	−4	−10	−10	−6	−2	−3	11	−10	4
$a_7a_2a_1^2$	−4	−8	41	17	22	19	15	4	7	−9	−11	−15
$a_7a_1^4$	.	.	−12	−5	−6	−5	−4	−1	.	3	6	5
a_6a_5	18	3	−24	3	−27	3	33	11	−11	4	8	8
$a_6a_4a_1$	3	6	6	−24	18	−12	−18	−12	−6	7	−14	14
$a_6a_3a_2$	−3	−6	30	−12	3	−21	−15	−5	10	−17	−4	8
$a_6a_3a_1^2$	−7	−4	−11	23	−5	22	19	5	7	6	1	−21
$a_6a_2^2a_1$	4	3	−18	16	−1	23	17	5	−7	6	14	−12
$a_6a_2a_1^3$	.	.	7	−22	1	−24	−19	−5		−3	−6	20
$a_6a_1^5$	.	.	.	5	.	5	4	1	.	.	.	−5
$a_5^2a_1$	−13	2	9	2	12	7	−23	−6	11	−4	−8	−8
$a_5a_4a_2$	5	−2	−18	14	11	1	−11	−13	−6	7	8	−4
$a_5a_4a_1^2$	5	−2	2	−1	−15	−5	22	13	−5	−3	11	3
$a_5a_3^2$	.	−6	6	6	−3	6	−9	−3	−1	2	7	−17
$a_5a_3a_2a_1$	−1	8	−7	−18	2	−11	41	12	−2	6	−18	30
$a_5a_3a_1^3$	.	.	4	2	2	3	−23	−6	.	−3	2	−4
$a_5a_2^3$	.	−3	6	−6	.	.	6	2	3	−6	.	6
$a_5a_2^2a_1^2$	.	.	−3	12	−1	3	−32	−9	.	3	2	−18
$a_5a_2a_1^4$	.	.	.	−3	.	−1	23	6	.	.	.	5
$a_5a_1^6$	.	.	.	.	.	.	−4	−1	.	.	.	.
$a_4^2a_3$	−1	5	−2	−7	−5	7	7	−7	1	−1	−6	10
$a_4^2a_2a_1$	−2	−1	8	3	−2	−4	−8	14	5	−6	−2	−6
$a_4^2a_1^3$	.	.	−3	−1	3	2	1	−7	.	3	−3	1
$a_4a_3^2$	1	−2	−1	4	5	−7	−4	7	−1	.	5	−3
$a_4a_3a_2^2$	2	1	−2	2	−1	3	−7	7	−3	6	2	−12
$a_4a_3a_2a_1^2$	5	2	1	−4	−3	6	9	−21	.	−3	4	6
$a_4a_3a_1^4$	12	6	4	1	.	−2	−1	7	.	.	.	−1
$a_4a_2^3a_1$	12	7	3	.	1	−3	5	−7	.	.	−2	6
$a_4a_2^2a_1^3$	31	18	11	2	3	1	−5	14	.	.	.	−2
$a_4a_2a_1^5$	80	50	35	11	10	5	1	−7	.	.	.	.
$a_4a_1^7$	210	140	105	42	35	21	7	1	.	.	.	.
$a_3^3a_2$	7	3	.	.	.	.	.	.	1	−2	−1	5
$a_3^3a_1^3$	17	6	3	.	.	.	.	.	2	1	−2	−1
$a_3^2a_2^2a_1$	36	19	8	.	2	.	.	.	5	2	1	−3
$a_3^2a_2a_1^4$	89	48	29	5	6	2	.	.	12	7	3	1
$a_3^2a_1^6$	220	130	90	27	20	10	2	.	30	20	10	5
$a_3a_2^3$	78	48	22	.	9	.	.	.	12	6	4	.
$a_3a_2^2a_1^3$	197	121	71	12	25	7	.	.	31	18	11	3
$a_3a_2a_1^5$	497	316	213	60	74	33	5	.	80	53	32	14
$a_3a_2^4a_1$	1260	840	611	223	225	123	33	3	210	150	95	51
$a_3a_1^8$	3220	2240	1708	728	686	420	148	25	560	420	280	168
$a_2^5a_1$	430	285	170	30	75	20	.	.	80	50	35	10
$a_2^4a_1^3$	1092	735	492	138	210	88	12	.	213	144	90	42
$a_2^3a_1^5$	2775	1920	1383	495	600	311	75	6	570	411	285	148
$a_2^2a_1^7$	7070	5040	3808	1589	1722	1008	324	48	1540	1162	826	483
$a_2a_1^9$	18060	13230	10332	4788	4914	3108	1188	243	4200	3276	2394	1512
a_1^{11}	46200	34650	27720	13860	13860	9240	3960	990	11550	9240	6930	4620

Table 2.11 (*cont.*)

$w = 11$ (v)	(7^2)	(5^21)	(641)	(731)	(821)	(91^2)	(65)	(74)	(83)	(92)	(10, 1)	(11)
a_{11}	11	11	22	22	22	11	−11	−11	−11	−11	−11	11
$a_{10}a_1$	−11	−6	−12	−12	−12	−1	11	11	11	11	1	−11
a_9a_2	7	−11	−22	−22	−4	−2	11	11	11	−7	11	−11
$a_9a_1^2$	2	6	12	12	3	1	−11	−11	−11	−2	−1	11
a_8a_3	−11	−11	−22	2	2	−11	11	11	−13	11	11	−11
$a_8a_2a_1$	4	17	34	10	8	3	−22	−22	2	−4	−12	22
$a_8a_1^3$	−2	−6	−12	−4	−3	−1	11	11	3	2	1	−11
a_7a_4	3	−11	6	6	−22	−11	11	−17	11	11	11	−11
$a_7a_3a_1$	8	17	6	3	10	12	−22	6	2	−22	−12	22
$a_7a_2^2$	.	11	8	8	4	2	−11	3	−11	7	−11	11
$a_7a_2a_1^2$	−6	−23	−18	−15	−11	−4	33	5	9	6	13	−33
$a_7a_1^4$	2	6	5	4	3	1	−11	−4	−3	−2	−1	11
a_6a_5	−11	19	8	−22	−22	−11	−19	11	11	11	11	−11
$a_6a_4a_1$	8	−13	.	6	34	12	8	6	−22	−22	−12	22
$a_6a_3a_2$	4	−8	14	20	2	13	8	−22	2	−4	−22	22
$a_6a_3a_1^2$	−10	7	−12	−15	−13	−13	3	5	9	24	13	−33
$a_6a_2^2a_1$	−4	2	−24	−18	−12	−5	3	19	9	−3	23	−33
$a_6a_2a_1^3$	8	−1	24	19	14	5	−14	−16	−12	−8	−14	44
$a_6a_1^5$	−2	.	−5	−4	−3	−1	5	4	3	2	1	−11
$a_5^2a_1$	11	−9	−13	17	17	6	19	−11	−11	−11	−6	11
$a_5a_4a_2$	−10	−8	−14	16	26	13	8	6	−22	−4	−22	22
$a_5a_4a_1^2$	−10	12	18	−18	−37	−13	−27	5	33	24	13	−33
$a_5a_3^2$	11	−4	7	−2	−2	11	4	−11	13	−11	−11	11
$a_5a_3a_2a_1$	−16	9	6	−33	−20	−28	−24	38	−6	30	46	−66
$a_5a_3a_1^3$	12	−6	−6	19	16	14	10	−16	−12	−26	−14	44
$a_5a_2^3$	.	−1	2	−8	−4	−2	1	−3	11	−7	11	−11
$a_5a_2^2a_1^2$	10	1	−3	33	23	9	9	−24	−18	−3	−36	66
$a_5a_2a_1^4$	−10	.	1	−23	−17	−6	−5	20	15	10	15	−55
$a_5a_1^6$	2	.	.	4	3	1	.	−4	−3	−2	−1	11
$a_4^2a_3$	−3	11	−6	−18	10	11	−11	17	1	−11	−11	11
$a_4^2a_2a_1$	13	2	6	12	−34	−14	3	−23	21	15	23	−33
$a_4^2a_1^3$	−1	−3	−5	−3	20	7	8	6	−18	−13	−7	22
$a_4a_3^2a_1$	−5	−13	3	27	−8	−23	18	−23	−15	33	23	−33
$a_4a_3a_2^2$	10	−3	6	.	−6	−15	3	−9	9	−3	33	−33
$a_4a_3a_2a_1^2$	−10	7	−12	−21	44	45	−12	36	−12	−60	−72	132
$a_4a_3a_1^4$	.	.	5	5	−19	−15	−5	−8	15	28	15	−55
$a_4a_2^3a_1$	−10	−1	2	−2	16	7	−4	12	−20	10	−34	44
$a_4a_2^2a_1^3$	10	.	−1	4	−37	−14	5	−16	30	11	50	−110
$a_4a_2a_1^5$	−2	.	.	−1	20	7	.	4	−18	−12	−16	66
$a_4a_1^7$	.	.	.	.	−3	−1	.	.	3	2	1	−11
$a_3^3a_2$	−5	4	−7	2	8	−8	−4	11	−13	5	11	−11
$a_3^3a_1^2$	5	2	3	−9	−5	12	−7	6	14	−19	−12	22
$a_3^2a_2^2a_1$	−5	−4	6	.	−17	24	9	−24	30	−12	−51	66
$a_3^2a_2a_1^3$	−5	.	−3	8	11	−32	5	−2	−30	47	50	−110
$a_3^2a_1^5$	1	.	.	−2	−1	8	.	2	3	−15	−8	33
$a_3a_2^3$	.	1	−2	2	−2	2	−1	3	−5	7	11	11
$a_3a_2^2a_1^5$	.	2	1	−4	9	−16	−5	12	−10	−7	70	−110
$a_3a_2^3a_1^4$	2	6	4	1	−6	20	.	−4	15	−21	−65	165
$a_3a_2a_1^6$	13	20	15	6	1	−8	.	.	−3	14	17	−77
$a_3a_1^8$	56	70	56	28	8	1	.	.	.	−2	−1	11
$a_2^5a_1$	.	11	5	.	.	.	1	−3	5	−7	9	−11
$a_2^4a_1^3$	6	30	19	4	.	.	3	1	−5	14	−30	55
$a_2^3a_1^5$	36	90	65	23	3	.	10	5	1	−7	27	−77
$a_2^2a_1^7$	155	280	217	98	23	2	35	21	7	1	−9	44
$a_2a_1^9$	576	882	714	372	117	19	126	84	36	9	1	−11
a_1^{11}	1980	2772	2310	1320	495	110	462	330	165	55	11	1

Table 2.12

$w=12$ (i)	(1^{12})	(21^{10})	(2^21^8)	(31^9)	(2^21^8)	(321^7)	(41^8)	(2^41^4)	(32^21^6)	(3^21^6)
a_{12}	1	−12	54	12	−112	−96	−12	105	252	42
$a_{11}a_1$	12	1	−10	−1	35	19	1	−50	−87	−9
$a_{10}a_2$	66	10	1	−2	−8	16	2	20	−32	−17
$a_{10}a_1^2$	132	21	2	1	.	−9	−1	.	27	9
a_9a_3	220	45	8	.	1	−3	3	−6	18	3
$a_9a_2a_1$	660	145	26	9	3	1	−3	.	−7	−1
$a_9a_1^3$	1320	300	54	28	6	3	1	.	.	.
a_8a_4	495	120	28	.	6	.	.	1	−4	2
$a_8a_3a_1$	1980	525	128	36	27	7	.	4	1	−2
$a_8a_2^2$	2970	810	201	72	42	15	.	6	2	1
$a_8a_2a_1^2$	5940	1665	418	189	87	38	8	12	5	2
$a_8a_1^4$	11880	3420	870	468	180	94	33	24	12	6
a_7a_5	792	210	56	.	15	.	.	4	.	.
$a_7a_4a_1$	3960	1170	336	84	93	21	.	24	5	.
$a_7a_3a_2$	7920	2460	736	252	207	70	.	52	17	6
$a_7a_3a_1^2$	15840	5040	1528	624	432	168	28	108	40	12
$a_7a_2^3a_1$	23760	7740	2392	1080	681	302	56	168	72	32
$a_7a_2a_1^3$	47520	15840	4968	2520	1422	703	204	348	166	78
$a_7a_1^5$	95040	32400	10320	5760	2970	1620	600	720	380	200
a_6^2	924	252	70	.	20	.	.	6	.	.
$a_6a_5a_1$	5544	1722	532	126	165	35	.	52	10	.
$a_6a_4a_2$	13860	4620	1526	504	498	161	.	160	50	15
$a_6a_4a_1^2$	27720	9450	3164	1218	1041	378	56	336	115	30
$a_6a_3^2$	18480	6300	2128	756	707	252	.	228	81	30
$a_6a_3a_2a_1$	55440	19740	6888	3024	2337	1008	168	760	320	132
$a_6a_3a_1^3$	110880	40320	14280	6888	4890	2289	588	1596	722	306
$a_6a_2^3$	83160	30240	10752	5040	3696	1722	336	1266	552	262
$a_6a_2^2a_1^2$	166320	61740	22288	11340	7737	3878	1064	2532	1241	608
$a_6a_2a_1^4$	332640	126000	46200	25200	16200	8680	2940	5316	2780	1440
$a_6a_1^6$	665280	257040	95760	55440	33930	19320	7560	11160	6210	3440
$a_5^2a_2$	16632	5670	1932	630	660	210	.	228	70	20
$a_5^2a_1^2$	33264	11592	4004	1512	1380	490	70	480	160	40
$a_5a_4a_3$	27720	9870	3528	1260	1266	455	.	456	165	60
$a_5a_4a_2a_1$	83160	30870	11396	4914	4188	1771	280	1536	635	250
$a_5a_4a_1^3$	166320	63000	23604	11088	8766	3983	966	3240	1420	570
$a_5a_3^2a_1$	110880	42000	15848	7056	5952	2632	420	2220	976	420
$a_5a_3a_2^2$	166320	64260	24696	11592	9417	4410	840	3552	1652	776
$a_5a_3a_2a_1^2$	332640	131040	51128	25704	19722	9793	2548	7500	3670	1762
$a_5a_3a_1^4$	665280	267120	105840	56448	41310	21644	6888	15840	8132	4056
$a_5a_2^3a_1$	498960	200340	79632	41580	31221	16212	4536	12012	6177	3192
$a_5a_2^2a_1^3$	997920	408240	164808	90720	65412	35658	11844	25380	13656	7308
$a_5a_2a_1^5$	1995840	831600	341040	196560	137070	78120	29400	53640	30130	16800
$a_5a_1^7$	3991680	1693440	705600	423360	287280	170520	70560	113400	66360	38640
a_4^3	34650	12600	4620	1680	1710	630	.	639	240	90
$a_4^2a_3a_1$	138600	53550	20720	9240	8040	3570	560	3132	1390	600
$a_4^2a_2^2$	207900	81900	32270	15120	12720	5950	1120	5022	2340	1090
$a_4^2a_2a_1^2$	415800	166950	66780	33390	26640	13160	3360	10620	5180	2460
$a_4^2a_1^4$	831600	3402000	138180	73080	55800	28980	9030	22464	11440	5620
$a_4a_3^2a_2$	277200	111300	44800	21420	18075	8680	1680	7308	3525	1710
$a_4a_3^2a_1^2$	554400	226800	92680	47040	37860	19110	4900	15468	7780	3840
$a_4a_3a_2^2a_1^2$	831600	346500	144200	75600	59925	31360	8680	24864	12980	6760
$a_4a_3a_2a_1^3$	1663200	705600	298200	163800	125550	68495	22260	52668	28530	15270
$a_4a_3a_1^5$	3326400	1436400	616560	352800	263070	149100	54600	111600	62600	34600
$a_4a_2^4$	1247400	529200	224280	120960	94860	51240	15120	40002	21600	11760
$a_4a_2^3a_1^2$	2494800	1077300	463680	260820	198765	111510	37800	84780	47385	26460
$a_4a_2^2a_1^4$	4989600	2192400	958440	559440	416520	241920	91140	179748	103780	59640
$a_4a_2a_1^6$	9979200	4460400	1980720	1194480	872910	523320	214200	381240	226950	134400
$a_4a_1^8$	19958400	9072000	4092480	2540160	1829520	1128960	493920	808920	495600	302400
a_3^4	369600	151200	62160	30240	25680	12600	2520	10656	5280	2640
$a_3^3a_2a_1$	1108800	470400	199920	105840	85140	45150	12600	36336	19320	10260
$a_3^3a_1^3$	2217600	957600	413280	228480	178380	98280	31920	77040	42360	23040
$a_3^2a_2^3$	1663200	718200	310800	168840	134760	73500	21840	58536	32040	17600
$a_3^2a_2^2a_1^2$	3326400	1461600	642320	362880	282360	153960	54040	124176	70120	39520
$a_3^2a_2a_1^4$	6652800	2973600	1327200	776160	591660	344960	129360	263520	153220	88560
$a_3^2a_1^6$	13305600	6048000	2741760	1653120	1239840	744240	302400	559440	334320	198400
$a_3a_2^4a_1$	4989600	2230200	997920	574560	446040	257880	90720	200328	115800	67200
$a_3a_2^3a_1^3$	9979200	4536000	2061360	1224720	936540	556290	214200	425376	252600	149940
$a_3a_2^2a_1^5$	19958400	9223200	4257120	2600640	1962540	1197000	495600	903600	550240	334320
$a_3a_2a_1^7$	39916800	18748800	8780760	5503680	4112640	2569560	1128960	1920240	1197000	744240
$a_3a_1^9$	79833600	38102400	18144000	11612160	8618400	5303680	2540160	4082400	2600640	1653120
a_2^6	7484400	3402000	1549800	907200	707400	415800	151200	323460	190800	113400
$a_2^5a_1^2$	14968800	6917400	3200400	1927800	1482300	894600	352800	687240	415500	252000
$a_2^4a_1^4$	29937600	14061600	6607440	4082400	3106080	1920240	808920	1460736	903600	559440
$a_2^3a_1^6$	59875200	28576800	13638240	8618400	6508620	4112640	1829520	3106080	1962540	1239840
$a_2^2a_1^8$	119750400	58060800	28143360	18144000	13638240	8789760	4092480	6607440	4257120	2741760
$a_2a_1^{10}$	239500800	117936000	58060800	38102400	28576800	18748800	9072000	14061600	9223200	6048000
a_1^{12}	479001600	239500800	119750400	79833600	59875200	39916800	19958400	29937600	19958400	13305600

Table 2.12 (*cont.*)

$w=12$ (ii)	(421^6)	(51^7)	$(2^6 1^5)$	$(32^2 1^5)$	$(3^3 21^4)$	$(42^2 1^4)$	(431^5)	(521^5)	(61^6)	(2^6)	$(32^4 1)$
a_{12}	84	12	−36	−240	−180	−180	−72	−72	−12	2	60
$a_{11}a_1$	−18	−1	25	130	70	70	17	17	1	−2	−49
$a_{10}a_2$	−14	−2	−16	.	70	20	32	12	2	2	20
$a_{10}a_1^2$	8	1	.	−30	−45	−20	−17	−7	−1	.	9
a_9a_3	21	−3	9	−21	−36	45	18	18	3	−2	−6
$a_9a_2a_1$	23	3	.	14	13	−54	−22	−20	−3	.	−7
$a_9a_1^3$	−8	−1	.	.	.	20	8	7	1	.	.
a_8a_4	4	−4	−4	16	−4	−20	−8	24	4	2	−4
$a_8a_3a_1$	−1	4	.	−5	10	5	5	−27	−4	.	5
$a_8a_2^2$	−2	2	.	.	−6	12	.	−12	−2	.	.
$a_8a_2a_1^2$	1	−4									
$a_8a_1^4$	4	1	.	.	.	−6	−1	27	4	.	.
a_7a_5	.	.	.	.	.	.	.	−7	−1	.	.
$a_7a_4a_1$	.	.	1	−5	5	5	−5	−5	5	−2	10
$a_7a_3a_2$	.	.	5	1	−3	−1	5	1	−5	.	−3
$a_7a_3a_1^2$	6	.	10	3	1	−2	−1	5	−5	.	.
$a_7a_2^2a_1$	13	.	20	7	2	1	−2	−1	5	.	.
$a_7a_2a_1^3$	46	7	30	12	5	2	1	−3	5	.	.
$a_7a_1^5$	135	36	120	60	27	12	7	3	1	−5	.
a_6^2	.	.	2	.	30	20	10	5	1	.	.
$a_6a_5a_1$	.	.	17	3	.	.	.	.	.	1	−6
$a_6a_4a_2$	.	.	50	15	4	.	.	.	.	6	1
$a_6a_4a_1^2$	15	.	105	34	8	4	.	.	.	15	4
$a_6a_3^2$	.	.	70	24	9	.	.	.	.	30	9
$a_6a_3a_2a_1$	51	.	230	93	37	14	5	.	.	20	6
$a_6a_3a_1^3$	174	21	480	208	84	47	15	5	.	60	22
$a_6a_2^3$	108	.	360	159	72	30	16	.	.	120	48
$a_6a_2^2a_1^2$	331	42	750	354	167	90	43	11	.	90	36
$a_6a_2a_1^4$	910	196	1560	786	392	245	120	50	6	180	78
$a_6a_1^6$	2355	672	3240	1740	930	630	335	171	37	360	168
$a_5^2a_2$	.	.	81	24	6	.	.	.	.	720	360
$a_5^2a_1^2$	20	.	172	54	12	6	.	.	.	30	9
$a_5a_4a_3$	.	.	165	60	22	.	.	.	.	66	20
$a_5a_4a_2a_1$	95	.	565	228	86	.	.	.	.	60	22
$a_5a_4a_1^3$	320	35	1200	507	192	32	10	.	.	210	83
$a_5a_3^2a_1$	150	.	820	357	154	106	30	10	.	450	184
$a_5a_3a_2^2$	315	.	1320	606	279	53	20	.	.	300	128
$a_5a_3a_2a_1^2$	926	105	2800	1344	632	114	55	.	.	480	216
$a_5a_3a_1^4$	2489	476	5940	2976	1442	327	145	35	.	1020	476
$a_5a_2^3a_1$	1713	210	4500	2277	1153	872	390	155	15	2160	1048
$a_5a_2^2a_1^3$	4446	882	9540	5037	2640	618	321	75	.	1620	798
$a_5a_2a_1^5$	11085	2856	20220	11130	6070	1593	828	306	30	3420	1752
$a_5a_1^7$	26880	8232	42840	24570	14000	3970	2135	985	171	7200	3840
a_4^3	.	.	240	93	36	9660	5460	2856	672	15120	8400
$a_4^2a_3a_1$	210	.	1225	546	238	.	.	.	.	90	36
$a_4^2a_2^2$	440	.	1990	924	424	80	30	.	.	480	216
$a_4^2a_2a_1^2$	1280	140	4245	2046	954	172	80	.	.	795	368
$a_4^2a_1^4$	3420	630	9060	4524	2160	488	210	50	.	1710	817
$a_4a_3^2a_2$	690	.	2960	1434	699	1294	560	220	20	3690	1812
$a_4a_3^2a_1^2$	1950	210	6320	3174	1570	284	140	.	.	1200	584
$a_4a_3a_2^2a_1$	3575	420	10300	5361	2779	781	360	80	.	2580	1296
$a_4a_3a_2a_1^3$	9110	1715	22020	11841	6284	1466	755	170	.	4260	2210
$a_4a_3a_1^5$	22425	5460	47100	26130	14350	3711	1905	675	60	9180	4900
$a_4a_2^4$	6420	840	16800	9054	4900	9140	4800	2135	335	19800	10860
$a_4a_2^3a_1^2$	15975	3150	35940	19980	11085	2694	1500	360	.	7020	3768
$a_4a_2^2a_1^4$	38670	9660	76920	44088	25120	6678	3705	1305	120	15120	8352
$a_4a_2a_1^6$	91725	26880	164700	97200	56990	16201	9140	3970	630	32580	18504
$a_4a_1^8$	214200	70560	352800	214200	129360	38670	22425	11085	2355	70200	40080
a_3^4	1080	.	4440	2220	1128	468	240	.	.	1860	936
$a_3^3a_2a_1$	5400	630	13540	8289	4428	2328	1230	270	.	6660	3564
$a_3^3a_1^3$	13590	2520	33300	18300	9972	5820	3060	1050	90	14400	7920
$a_3^2a_2^3$	9030	1260	25470	13986	7728	4248	2390	570	.	11100	6114
$a_3^2a_2^2a_1^2$	23720	4620	54640	30858	17404	10428	5840	2020	180	24060	13592
$a_3^2a_2a_1^4$	56990	14000	117300	68040	39244	25120	14250	6070	930	52200	30220
$a_3^2a_1^6$	134400	38640	252000	149940	88560	59640	34600	16800	3440	113400	67200
$a_3a_2^4a_1$	40980	8400	89850	52020	30220	18504	10860	3840	360	40320	23376
$a_3a_2^3a_1^3$	97200	24570	193140	114657	68040	44088	26130	11130	1740	87660	52020
$a_3a_2^2a_1^5$	226950	66360	415500	252000	153220	103780	62600	30130	6210	190800	115800
$a_3a_2a_1^7$	523320	170520	894600	556290	344960	241920	149100	78120	19320	415800	257880
$a_3a_1^9$	1194480	423360	1927800	1224720	770160	559440	352800	196560	55440	907200	574560
a_2^6	70200	15120	148140	87660	52200	32580	19800	7200	720	67950	40320
$a_2^5a_1^2$	164700	42840	318930	193140	117300	76920	47100	20220	3240	148140	89850
$a_2^4a_1^4$	381240	113400	687240	425376	263520	179748	111600	53640	11160	323460	200328
$a_2^3a_1^6$	872910	287280	1482300	936540	591060	416520	263070	137070	33930	707400	446940
$a_2^2a_1^8$	1980720	705600	3200400	2061360	1327200	958440	616560	341040	95760	1549800	997920
$a_2a_1^{10}$	4460400	1693440	6917400	4536000	2973600	2192400	1436400	831600	257040	3402000	2230200
a_1^{12}	9979200	3991680	14968800	9979200	6652800	4989600	3326400	1995840	665280	7484400	4989600

Table 2.12 (*cont.*)

$w=12$ (iii)	$(3^2 2^2 1^2)$	$(42^2 1^2)$	$(3^2 1^3)$	(4321^2)	$(52^2 1^2)$	$(4^2 1^4)$	(531^4)	(621^6)	(71^6)	$(3^2 2^3)$	(42^4)	$(3^2 21)$
a_{12}	180	120	40	240	120	30	60	60	12	−24	−12	−48
$a_{11}a_1$	−114	−76	−18	−108	−54	−8	−16	−16	−1	24	12	37
$a_{10}a_2$	−45	10	−30	−100	−10	−15	−30	−10	−2	−6	−8	28
$a_{10}a_1^2$	54	16	18	68	14	8	16	6	1	−9	−2	−27
a_9a_3	63	−39	17	−33	−30	−21	−15	−15	−3	−3	12	−42
$a_9a_2a_1$	−30	39	−6	73	37	23	19	17	3	9	−4	19
$a_9a_1^3$	.	−16	.	−32	−14	−8	−7	−6	−1	.	2	.
a_8a_4	−28	24	8	48	−40	6	−20	−20	−4	8	−12	16
$a_8a_3a_1$	−5	−5	−11	−27	52	−3	23	23	4	−5	.	13
$a_8a_2^2$	9	−18	6	−12	18	−1	14	10	2	−2	4	−12
$a_8a_2a_1^2$	.	9	.	11	−51	1	−27	−23	−4	.	−2	.
$a_8a_1^4$	.	.	.	.	14	.	7	6	1	.	.	.
a_7a_5	−5	−15	−5	5	20	5	10	−25	−5	−11	12	13
$a_7a_4a_1$	9	3	3	−17	−4	−5	−6	29	5	3	.	−18
$a_7a_3a_2$	−4	8	−1	7	−23	1	−4	25	5	2	−4	7
$a_7a_3a_1^2$	.	−4	.	8	4	2	3	−29	−5	.	2	.
$a_7a_2^2a_1$	.	.	.	−5	15	−1	2	−27	−5	.	.	.
$a_7a_2a_1^3$	.	.	.	.	−5	.	−1	29	5	.	.	.
$a_7a_1^5$	.	.	.	.	.	.	.	−6	−1	.	.	.
a_6^2	9	6	−2	−12	−6	3	6	6	−6	6	−6	−12
$a_6a_5a_1$	−4	−1	2	7	1	−3	−6	−1	6	−1	.	7
$a_6a_4a_2$	1	−2	−2	4	2	−3	2	−6	6	−2	4	4
$a_6a_4a_1^2$	2	1	.	−3	−1	3	2	1	−6	.	−2	.
$a_6a_3^2$	2	.	1	−3	3	3	−3	−3	3	.	.	−3
$a_6a_3a_2a_1$	8	3	3	1	−3	−3	4	7	−12	.	.	.
$a_6a_3a_1^3$	18	10	6	3	1	.	−2	−1	6	.	.	.
$a_6a_2^3$	15	6	6	3	.	1	−2	2	−2	.	.	.
$a_6a_2^2a_1^2$	34	18	15	8	2	2	1	−4	9	.	.	.
$a_6a_2a_1^4$	78	48	36	22	9	6	4	1	−6	.	.	.
$a_1a_1^5$	180	120	90	60	30	20	15	6	1	.	.	.
$a_5^2a_2$	2	.	.	.	.	.	.	.	.	1	−2	−3
$a_5^2a_1^2$	4	2	.	.	.	.	.	.	.	2	1	.
$a_5a_4a_3$	8	.	3	.	.	.	.	.	.	3	.	1
$a_5a_4a_2a_1$	30	11	9	3	.	.	.	.	.	11	4	3
$a_5a_4a_1^3$	66	36	18	9	3	.	.	.	.	24	13	6
$a_5a_3^2a_1$	54	18	24	7	.	.	.	.	.	18	6	7
$a_5a_3a_2^2$	96	39	42	18	.	4	.	.	.	31	12	12
$a_5a_3a_2a_1^2$	216	110	96	47	11	8	4	.	.	68	34	27
$a_5a_3a_1^4$	488	291	216	124	48	24	16	4	.	150	88	60
$a_5a_2^3a_1$	390	207	189	103	24	30	13	.	.	117	60	51
$a_5a_2^2a_1^3$	886	528	444	264	96	78	48	9	.	258	150	117
$a_5a_2a_1^5$	2020	1305	1050	675	306	220	155	50	5	570	360	270
$a_5a_1^7$	4620	3150	2520	1715	882	630	476	196	36	1260	840	630
a_4^3	15	.	6	.	.	.	.	.	.	6	.	3
$a_4^2a_3a_1$	96	31	42	12	.	.	.	.	.	39	12	18
$a_4^2a_2^2$	168	68	72	30	.	6	.	.	.	68	28	30
$a_4^2a_2a_1^2$	376	189	162	78	18	12	6	.	.	153	76	66
$a_4^2a_1^4$	844	498	360	204	78	36	24	6	.	344	201	144
$a_4a_3^2a_2$	286	117	141	58	.	12	.	.	.	117	48	58
$a_4a_3^2a_1^2$	644	314	318	147	31	24	12	.	.	264	126	132
$a_4a_3a_2a_1^3$	1140	599	582	305	68	80	34	.	.	467	244	237
$a_4a_3a_1^5$	2578	1506	1332	763	264	204	124	22	.	1056	610	540
$a_4a_2^3a_1$	5840	3705	3060	1905	828	560	390	120	10	2390	1500	1230
$a_4a_2^2a_1^4$	2026	1116	1080	610	150	201	88	.	.	828	456	438
$a_4a_2a_1^6$	4592	2754	2505	1506	528	498	291	48	.	1875	1116	1008
$a_4a_1^8$	10428	6678	5820	3711	1593	1294	872	245	20	4248	2694	2328
$a_3^3a_2$	23720	15975	13590	9110	4446	3420	2489	910	135	9630	6420	5400
$a_3^3a_1^2$	54040	37800	31920	22260	11844	9030	6888	2940	600	21840	15120	12600
$a_3^2a_2^2$	480	204	256	108	.	24	.	.	.	204	90	108
$a_3^2a_2a_1$	1914	1008	1038	540	117	144	60	.	.	828	438	451
$a_3^2a_1^3$	4332	2505	2364	1332	444	360	216	36	.	1884	1080	1038
$a_3^2a_2^3$	3386	1875	1884	1056	258	344	150	.	.	1486	828	828
$a_3^2a_2^2a_1^2$	7672	4592	4332	2578	886	844	488	78	.	3386	2026	1914
$a_3^2a_2a_1^4$	17404	11085	9972	6284	2640	2160	1442	392	30	7728	4900	4428
$a_3^2a_1^6$	39520	26460	23040	15270	7308	5620	4056	1440	200	17660	11760	10260
$a_3a_2^4$	13592	8352	7920	4900	1752	1812	1048	168	.	6114	3768	3564
$a_3a_2^3a_1^2$	30858	19986	18300	11841	5037	4524	2976	786	60	13986	9054	8289
$a_3a_2^2a_1^4$	70120	47385	42360	28530	13656	11440	8132	2780	380	32040	21600	19320
$a_3a_2a_1^6$	159460	111510	98280	68495	35658	28980	21644	8680	1620	73500	51240	45150
$a_3a_1^8$	362880	260820	228480	163800	90720	73080	56448	25200	5760	168840	120960	105840
a_2^6	24060	15120	14400	9180	3420	3690	2160	360	.	11100	7020	6660
$a_2^4a_1^2$	54640	35940	33300	22020	9540	9060	5940	1560	120	25470	16800	15540
$a_2^3a_1^4$	124176	84780	77040	52668	23380	22464	15840	5316	720	58536	40002	36336
$a_2^2a_1^6$	282360	198765	178380	125550	56412	55800	41310	16200	2970	134760	94860	85140
$a_2a_1^8$	642320	463680	413280	298200	164808	138180	105840	46200	10320	310800	224280	199920
a_1^{10}	1461600	1077300	957600	705600	408240	340200	267120	126000	32400	718200	529200	479400
	3326400	2494800	2217600	1663200	997920	831600	665280	332640	95040	1663200	1247400	1108800

Table 2.12 (*cont.*)

$w=12$ (iv)	(432^21)	(52^21)	(43^21^3)	(4^321^5)	(5321^3)	(62^21^3)	(541^3)	(631^3)	(721^3)	(81^4)	(3^4)	(43^22)	(4^32^2)
a_{12}	−144	−48	−72	−72	−144	−72	−48	−48	−48	−12	3	36	18
$a_{11}a_1$	111	37	39	39	78	39	15	15	15	1	−3	−36	−18
$a_{10}a_2$	24	−12	62	32	64	2	28	28	8	2	−3	−16	2
$a_{10}a_1^2$	−51	−7	−39	−24	−48	−9	−15	−15	−5	−1	3	26	8
a_9a_3	9	21	−9	45	9	18	39	12	12	3	6	18	−18
$a_9a_2a_1$	−36	−16	−20	−62	−43	−23	−43	−16	−14	−3	−3	−2	16
$a_9a_1^3$	24	7	12	24	21	9	15	6	5	1	.	−8	−8
a_8a_4	−16	16	−40	−40	48	24	16	16	16	4	−3	−4	22
$a_8a_3a_1$	16	−26	34	16	−63	−33	−22	−19	−19	−4	−3	−14	−4
$a_8a_2^2$	24	−4	−6	16	−32	−10	−12	−12	−8	−2	3	.	−14
$a_8a_2a_1^2$	−17	23	−5	−10	75	32	26	23	19	4	.	8	8
$a_8a_1^4$	.	−7	.	.	−21	−9	−7	−6	−5	−1	.	.	.
a_7a_5	4	−22	2	2	−31	37	−22	13	13	12	−3	−1	−18
$a_7a_4a_1$	3	3	15	15	21	−42	11	−24	−24	−5	6	5	−4
$a_7a_3a_2$	−19	19	−4	−7	39	−27	3	−26	−20	−5	−3	5	16
$a_7a_3a_1^2$	.	−2	−9	−6	−15	42	−5	27	24	5	.	2	−4
$a_7a_2^2a_1$	5	−15	5	4	−23	33	−1	28	22	5	.	−5	−2
$a_7a_2a_1^3$	.	5	.	.	9	−41	1	−29	−24	−5	.	.	.
$a_7a_1^5$	.	.	.	.	.	9	.	6	5	1	.	.	.
a_6^2	.	12	18	.	.	−18	−12	−12	24	6	3	.	9
$a_6a_5a_1$	−1	−1	−20	1	7	2	19	14	−28	−13	−3	1	.
$a_6a_4a_2$	−8	−4	2	8	−16	22	4	4	−24	−6	−3	8	−12
$a_6a_4a_1^2$	6	2	3	−9	−3	−3	−8	−3	29	6	.	−7	8
$a_6a_3^2$	9	−9	.	−9	9	9	−3	6	−12	−3	3	−9	.
$a_6a_3a_2a_1$	−3	9	−1	11	−11	−25	2	−11	53	12	.	5	−4
$a_6a_3a_1^3$	.	−3	.	.	6	3	2	3	−29	−6	.	.	.
$a_6a_2^3$	.	.	.	−4	8	−8	.	.	8	2	.	.	2
$a_6a_2^2a_1^2$	.	.	.	.	−4	16	−1	3	−41	−9	.	.	.
$a_6a_2a_1^4$	.	.	.	.	.	−4	.	−1	29	6	.	.	.
$a_6a_1^6$	.	.	.	.	.	.	.	.	−5	−1	.	.	.
$a_5^2a_2$	6	2	3	−7	−4	−2	7	−3	7	−7	3	−9	8
$a_5^2a_1^2$	−4	−1	2	4	3	1	−7	−2	−1	7	.	4	−4
$a_5a_4a_3$	−3	3	−1	5	−2	−7	−5	7	7	−7	−3	11	−4
$a_5a_4a_2a_1$	1	−3	−2	−1	8	3	−2	−4	−8	14	.	−1	.
$a_5a_4a_1^3$	3	1	.	.	−3	−1	3	2	1	−7	.	.	.
$a_5a_3^2a_1$	2	.	1	−2	−1	4	5	−7	−4	7	.	−2	4
$a_5a_3a_2^2$	5	.	2	1	−2	2	−1	3	−7	7	.	.	−2
$a_5a_3a_2a_1^2$	13	3	5	2	1	−4	−3	6	9	−21	.	.	.
$a_5a_3a_1^4$	34	13	12	6	4	1	.	−2	−1	7	.	.	.
$a_5a_2^3a_1$	27	6	12	7	3	.	1	−3	5	−7	.	.	.
$a_4a_3^2a_2^2$	68	24	31	18	11	2	3	1	−5	14	.	.	.
$a_4a_3^2a_1^3$	170	75	80	50	35	11	10	5	1	−7	.	.	.
$a_4a_1^7$	420	210	210	140	105	42	35	21	7	1	.	.	.
a_4^3	.	.	.	.	.	.	.	.	.	.	1	−4	2
$a_4^2a_3a_1$	5	.	2	.	.	.	.	.	.	.	4	1	−2
$a_4^2a_2^2$	12	.	4	2	.	.	.	.	.	.	6	2	1
$a_4^2a_2a_1^3$	31	7	10	4	2	.	.	.	.	.	12	5	2
$a_4^2a_1^5$	80	30	24	12	8	2	.	.	.	.	24	12	6
$a_4a_3^3a_2$	24	.	12	5	.	.	.	.	.	.	12	5	2
$a_4a_3^2a_1^3$	60	12	29	10	5	.	.	.	.	.	28	12	4
$a_4a_3a_2^3a_1$	123	27	60	31	13	.	3	.	.	.	48	24	12
$a_4a_3a_2a_1^4$	305	103	147	78	47	8	9	3	.	.	108	58	30
$a_4a_3a_1^5$	755	321	360	210	145	43	30	15	3	.	240	140	80
$a_4a_2^4$	244	60	126	76	34	.	13	.	.	.	90	48	28
$a_4a_2^3a_1^3$	590	207	314	180	110	18	36	10	.	.	204	117	68
$a_4a_2^2a_1^5$	1466	618	781	488	327	90	106	47	7	.	468	284	172
$a_4a_2a_1^7$	3575	1713	1950	1280	926	331	320	174	46	4	1080	690	440
$a_4a_1^8$	8080	4530	4900	3360	2548	1004	966	588	204	33	2520	1680	1120
a_3^4	48	.	28	12	.	.	.	.	.	.	24	12	6
$a_3^3a_2a_1$	237	51	132	66	27	.	6	.	.	.	108	58	30
$a_3^3a_1^3$	582	189	318	162	96	15	18	6	.	.	256	141	72
$a_3^2a_2^3$	407	117	264	153	68	.	24	.	.	.	204	117	68
$a_3^2a_2^2a_1^2$	1140	390	644	376	216	34	66	18	.	.	480	286	168
$a_3^2a_2a_1^4$	2779	1153	1570	954	632	107	192	84	12	.	1128	699	424
$a_3^2a_1^6$	6760	3192	3840	2460	1762	608	570	306	78	6	2640	1710	1090
$a_3a_2^4a_1$	2210	798	1296	817	476	78	184	48	.	.	936	584	368
$a_3a_2^3a_1^3$	5361	2277	3174	2046	1344	354	507	208	27	.	2220	1434	924
$a_3a_2^2a_1^5$	12980	6177	7780	5180	3670	1241	1420	722	166	12	5280	3525	2340
$a_3a_2a_1^7$	31360	16212	19110	13160	9703	3878	3983	2289	703	94	12600	8680	5950
$a_3a_1^9$	75600	41580	47040	33390	25704	11340	11088	6888	2520	468	30240	21420	15120
a_2^6	4260	1620	2580	1710	1020	180	450	120	.	.	1860	1200	795
$a_2^5a_1^2$	10300	4300	6320	4245	2800	750	1200	480	60	.	4440	2960	1990
$a_2^4a_1^4$	24864	12012	15468	10620	7500	2532	3240	1596	348	24	10656	7308	5022
$a_2^3a_1^6$	59925	31221	37860	26640	19722	7737	8766	4890	1422	180	25680	18075	12720
$a_2^2a_1^8$	144200	79632	92680	66780	51128	22288	23604	14280	4968	870	62160	44800	32270
$a_2a_1^{10}$	346500	200340	226800	166950	131040	61740	63000	40320	15840	3420	151200	111300	81900
a_1^{12}	831600	498960	554400	415800	332640	166320	166320	110880	47520	11880	369600	277200	207900

Table 2.12 (*cont.*)

$w=12$ (v)	(532^2)	(62^2)	(4^231)	(53^21)	(5421)	(6321)	(72^21)	(5^21^2)	(641^2)	(731^2)	(821^2)	(91^2)	(4^2)	(543)
a_{12}	36	12	36	36	72	72	36	18	36	36	36	12	−4	−24
$a_{11}a_1$	−36	−12	−25	−25	−50	−50	−25	−7	−14	−14	−14	−1	4	24
$a_{10}a_2$	4	8	−36	−36	−32	−32	4	−13	−26	−26	−6	−2	4	24
$a_{10}a_{11}$	16	2	25	25	30	30	5	7	14	14	4	1	−4	−24
a_9a_3	−9	−12	−9	18	−45	9	−9	−18	−36	−9	−9	−3	4	−3
$a_9a_2a_1$	5	4	34	7	73	19	12	20	40	13	11	3	−8	−21
$a_9a_1^3$	−7	−2	−16	−7	−30	−12	−5	−7	−14	−5	−4	−1	4	15
a_8a_4	−20	4	44	−4	−8	−40	−20	−18	−4	−4	−4	−12	−12	−8
$a_8a_3a_1$	29	8	−34	−1	31	33	18	25	18	15	15	4	8	11
$a_8a_2^2$	4	.	−4	20	.	16	4	13	10	10	6	2	4	−8
$a_8a_2a_1^2$	−21	−6	9	−24	−39	−41	−17	−27	−22	−19	−15	−4	−4	13
$a_8a_1^4$	7	2	.	7	14	12	5	7	6	5	4	1	.	−7
a_7a_5	34	−12	−1	−1	33	−37	−1	17	−1	−1	−36	−12	4	−11
$a_7a_4a_1$	−14	8	−26	−6	−19	55	24	−10	11	11	18	13	8	19
$a_7a_3a_2$	−23	4	10	−17	−28	30	12	−4	27	21	15	5	−8	14
$a_7a_3a_1^2$	4	−10	16	11	16	−49	−23	3	−25	−22	−19	−5	−4	−22
$a_7a_2^2a_1$	19	−4	−9	8	18	−42	−16	2	−29	−23	−17	−5	4	−6
$a_7a_2a_1^3$	−7	8	.	−4	−8	53	22	−1	29	24	19	5	.	7
$a_7a_1^5$	.	−2	.	.	.	−12	−5	.	−6	−5	−4	−1	.	.
a_6^2	−18	6	−18	.	.	36	−18	9	18	−18	−18	−6	2	12
$a_6a_5a_1$	2	.	26	5	−25	−27	26	−16	−27	15	50	13	−8	−13
$a_6a_4a_2$	16	−12	−8	4	16	−24	16	−5	−18	30	10	14	8	−16
$a_6a_4a_1^2$	−2	2	1	2	7	5	−29	9	17	−25	−22	−14	−4	5
$a_6a_3^2$	9	.	9	−9	9	−27	9	.	.	9	9	3	−4	3
$a_6a_3a_2a_1$	−11	8	−10	14	−10	56	−42	1	5	−49	−41	−12	8	−4
$a_6a_3a_1^3$	3	.	.	−7	−4	−11	28	−2	−3	27	23	6	.	7
$a_6a_2^3$	−4	4	4	−8	.	8	−4	−1	4	−10	−6	−2	−4	8
$a_6a_2^2a_1^2$	2	−8	.	4	3	−25	33	1	−3	42	32	9	.	−7
$a_6a_2a_1^4$	.	2	.	.	.	7	−27	.	1	−29	−23	−6	.	.
$a_6a_1^6$	.	.	.	.	.	.	5	.	.	5	4	1	.	.
$a_5^2a_2$	−4	2	1	1	−3	17	−19	−7	−4	−4	21	7	−4	11
$a_5^2a_1^2$	1	−1	−8	−5	10	1	2	7	9	3	−27	−7	6	1
$a_5a_4a_3$	−11	8	−10	11	8	−4	−6	1	5	−22	13	15	8	−14
$a_5a_4a_2a_1$	3	.	10	−10	−10	−10	18	10	7	16	−39	−30	−8	8
$a_5a_4a_1^3$	−1	.	.	5	−2	2	−1	−7	−8	−5	26	15	.	−5
$a_5a_3^2a_1$	2	−8	.	.	−10	14	8	−5	2	11	−24	−7	−4	11
$a_5a_3a_2^2$	4	−4	−1	2	3	−11	19	1	−2	4	−21	−7	4	−11
$a_5a_3a_2a_1^2$	−2	8	.	−1	8	−11	−23	3	−3	−15	75	21	.	−2
$a_5a_3a_1^4$	.	−2	.	.	.	4	2	.	2	3	−27	−7	.	.
$a_5a_2^3a_1$	.	.	.	.	−3	9	−15	−1	2	−2	23	7	.	3
$a_5a_2^2a_1^3$	.	.	.	.	.	−3	15	.	−1	4	−51	−14	.	.
$a_5a_2a_1^5$	.	.	.	.	.	.	−3	.	.	−1	27	7	.	.
$a_5a_1^7$	.	.	.	.	.	.	.	.	.	.	−4	−1	.	.
$a_4^2a_3$	4	−4	4	−4	−8	8	−4	6	−4	4	−4	4	−4	8
$a_4^2a_2a_1$	−1	4	−1	.	10	−10	−9	−8	1	10	9	−10	4	−10
$a_4^2a_1^3$	−2	2	−2	4	.	−4	−2	−4	8	−4	8	−8	2	−4
$a_4^2a_3a_1^2$	1	−4	.	−2	−1	11	4	4	−9	−6	−10	24	.	5
$a_4^2a_1^4$	4	1	.	.	−3	−1	.	3	2	1	−8	.	.	.
$a_4a_3^2a_2$	.	.	1	−2	−1	5	−5	4	−7	2	8	−8	−4	11
$a_4a_3^2a_1^2$	2	.	2	1	−2	−1	5	2	3	−9	−5	12	.	−1
$a_4a_3a_2^2a_1$	5	.	5	2	1	−3	5	−4	6	.	−17	24	.	−3
$a_4a_3a_2a_1^3$	18	3	12	7	3	1	−5	.	−3	8	11	−32	.	.
$a_4a_3a_1^5$	55	16	30	20	10	5	1	.	.	−2	−1	8	.	.
$a_4a_2^5$	12	.	12	6	4	.	.	1	−2	2	−2	2	.	.
$a_4a_2^3a_1^3$	39	6	31	18	11	3	.	2	1	−4	9	−16	.	.
$a_4a_2^2a_1^4$	114	30	80	53	32	14	2	6	4	1	−6	20	.	.
$a_4a_2a_1^6$	315	108	210	150	95	51	13	20	15	6	1	−8	.	.
$a_4a_1^8$	840	336	560	420	280	168	56	70	56	28	8	1	.	.
a_3^4	.	.	4	.	.	.	.	.	.	.	.	.	1	−3
$a_3^3a_2a_1$	12	.	18	7	3	.	.	.	.	.	.	.	3	1
$a_3^3a_1^3$	42	6	42	24	9	3	.	.	.	.	.	.	6	3
$a_3^2a_2^3$	31	.	39	18	11	.	.	2	.	.	.	.	6	3
$a_3^2a_2^2a_1^2$	96	15	96	54	30	8	.	4	2	.	.	.	15	8
$a_3^2a_2a_1^4$	279	72	238	154	86	37	5	12	8	2	.	.	30	22
$a_3^2a_1^6$	776	262	600	420	250	132	32	40	30	12	2	.	90	60
$a_3a_2^4a_1$	216	36	216	128	83	22	.	20	9	.	.	.	36	22
$a_3a_2^3a_1^3$	606	150	546	357	228	93	12	54	34	7	.	.	93	60
$a_3a_2^2a_1^5$	1652	552	1390	976	635	320	72	160	115	40	5	.	240	165
$a_3a_2a_1^7$	4410	1722	3570	2632	1771	1008	302	490	378	168	38	3	630	455
$a_3a_1^9$	11592	5040	9240	7056	4914	3024	1080	1512	1218	624	189	28	1680	1260
a_2^6	480	90	480	300	210	60	.	66	30	.	.	.	90	60
$a_2^4a_1^2$	1320	360	1225	820	565	230	30	172	105	20	.	.	240	165
$a_2^3a_1^4$	3552	1206	3132	2220	1536	700	168	480	336	108	12	.	639	456
$a_2^2a_1^6$	9417	3696	8040	5952	4188	2337	681	1380	1041	432	87	6	1710	1266
$a_2^3a_1^8$	24696	10752	20720	15848	11396	6888	2392	4004	3164	1528	418	54	4620	3528
$a_2a_1^{10}$	64260	30240	53550	42000	30870	19740	7740	11592	9450	5040	1665	300	12600	9870
a_1^{12}	166320	83160	138600	110880	85160	55440	23760	33264	27720	15840	5040	1320	34650	27720

Table 2.12 (*cont.*)

$w=12$ (vi)	(6^2)	(5^22)	(642)	(732)	(82^2)	(651)	(741)	(831)	(921)	$(10,1^2)$	(6^2)	(75)	(84)	(93)	(10, 2)	(11, 1)	(12)
a_{12}	−12	−12	−24	−24	−12	−24	−24	−24	−24	−12	6	12	12	12	12	12	−12
$a_{11}a_1$	12	12	24	24	12	13	13	13	13	1	−6	−12	−12	−12	−12	−1	12
$a_{10}a_2$	12	2	4	4	−8	24	24	24	4	2	−6	−12	−12	−12	−12	8	−12
$a_{10}a_1^2$	−12	−7	−14	−14	−2	−13	−13	−13	−3	−1	6	12	12	12	2	1	−12
a_9a_3	−15	12	24	−3	12	24	24	−3	−3	12	−6	12	−12	15	−12	−12	12
$a_9a_2a_1$	3	−14	−28	−1	−4	−37	−37	−10	−8	−3	12	24	24	−3	4	13	−24
$a_9a_1^3$	3	7	14	5	2	13	13	4	3	1	−6	−12	−12	−3	−2	−1	12
a_8a_4	12	12	−8	24	−4	24	−8	−8	24	12	−6	−12	20	−12	−12	−12	12
$a_8a_3a_1$	3	−24	−16	−21	−8	−37	−5	−2	−10	−13	12	24	−8	−3	24	13	−24
$a_8a_2^2$	−12	−2	12	−4	.	−24	−8	−8	−4	−2	6	12	−4	12	−8	12	−12
$a_8a_2a_1^2$	9	21	10	15	6	50	18	15	11	4	−18	−36	−4	−9	−6	−14	36
$a_8a_1^4$	−3	−7	−6	−5	−2	−13	−5	−4	−3	−1	6	12	4	3	2	1	−12
a_7a_5	12	−23	24	−11	12	−11	−11	24	24	12	−6	23	−12	−12	−12	−12	12
$a_7a_4a_1$	−24	11	−16	−13	−8	−2	2	−5	−37	−13	12	−11	−8	24	24	13	−24
$a_7a_3a_2$	3	21	−28	−8	−4	−13	−13	−21	−1	−14	12	−11	24	−3	4	24	−24
$a_7a_3a_1^2$	9	−4	30	21	10	15	11	15	13	14	−18	−1	−4	−9	−26	−14	36
$a_7a_2^2a_1$	9	−19	16	12	4	26	24	18	12	5	−18	−1	−20	−9	4	−25	36
$a_7a_2a_1^3$	−12	7	−24	−20	−8	−28	−24	−19	−14	−5	24	13	16	12	8	15	−48
$a_7a_1^5$	3	.	6	5	2	6	5	4	3	1	−6	−5	−4	−3	−2	−1	12
a_6^2	−12	6	−24	12	6	−24	12	12	12	6	15	−6	−6	−6	−6	−6	6
$a_6a_5a_1$	12	11	24	−13	−24	40	−2	−37	−37	−13	−24	−11	24	24	24	13	−24
$a_6a_4a_2$	12	−14	28	−28	12	24	−16	−16	−28	−14	−24	24	−8	24	4	24	−24
$a_6a_4a_1^2$	.	−4	−18	27	10	−27	11	18	40	14	18	−1	−4	−36	−26	−14	36
$a_6a_3^2$	6	−12	12	3	−12	12	−24	3	3	−12	−12	12	12	−15	12	12	−12
$a_6a_3a_2a_1$	−27	17	−24	30	16	−27	55	33	19	30	36	−37	−40	9	−32	−50	72
$a_6a_3a_1^3$	6	−3	4	−26	−12	14	−24	−10	−16	−15	−12	13	16	12	28	15	−48
$a_6a_2^3$	.	2	−12	4	.	.	8	8	4	2	6	−12	4	−12	8	−12	12
$a_6a_2^2a_1^2$	9	−2	22	−27	−10	2	−42	−33	−23	−9	−18	37	24	18	2	39	−72
$a_6a_2a_1^4$	−3	.	−6	25	10	−1	29	23	17	6	6	−25	−20	−15	−10	−16	60
$a_6a_1^6$	.	.	.	−5	−2	.	−5	−4	−3	−1	.	5	4	3	2	1	−12
$a_5^2a_4$	−12	3	−14	21	−2	11	11	−24	−14	−7	6	−23	12	12	2	12	−12
$a_5^2a_1^2$	.	−7	−5	−4	13	−16	−10	25	20	7	9	17	−18	−18	−13	−7	18
$a_5a_4a_3$	3	11	−16	14	−8	−13	19	11	−21	−24	12	−11	−8	−3	24	24	−24
$a_5a_4a_2a_1$	9	−3	16	−28	.	−25	−19	31	73	30	.	33	−8	−45	−32	−50	72
$a_5a_4a_1^3$	−3	7	4	3	−12	19	11	−22	−43	−15	−12	−22	16	39	28	15	−48
$a_5a_3^2a_1$	−9	1	4	−17	20	5	−6	−1	7	25	.	−1	−4	18	−36	−25	36
$a_5a_3a_2^2$	9	−4	16	−23	4	2	−14	29	5	16	−18	34	−20	−9	4	−36	36
$a_5a_3a_2a_1^2$	9	−4	−16	39	−32	7	21	−63	−43	−48	.	−31	48	9	64	78	−144
$a_5a_3a_1^4$	−3	.	2	−4	14	−6	−6	23	19	16	6	10	−20	−15	−30	−16	60
$a_5a_2^2a_1$	−9	2	−4	19	−4	−1	3	−26	−16	−7	12	−22	16	21	−12	37	−48
$a_5a_2a_1^3$	3	.	2	−23	18	1	−4	52	37	14	−6	20	−40	−30	−10	−54	120
$a_5a_2a_1^3$	.	.	.	5	−12	.	1	−27	−20	−7	.	−5	24	18	12	17	−72
$a_5a_1^7$	.	.	.	.	2	.	.	4	3	1	.	.	−4	−3	−2	−1	12
a_4^3	−4	−4	8	−8	4	−8	8	8	−8	−4	2	4	−12	4	4	4	−4
$a_4^2a_3a_1$	9	1	−8	10	−4	26	−26	−34	34	25	−18	−1	44	−9	−36	−25	36
$a_4^2a_2^2$	.	8	−12	16	−14	.	−4	−4	16	8	9	−18	22	−18	2	−18	18
$a_4^2a_2a_1^2$	−9	−7	8	−7	16	1	15	10	−62	−24	.	2	−40	45	32	39	−72
$a_4^2a_1^4$	3	.	−3	1	−1	−3	−5	−3	23	8	3	5	6	−21	−15	−8	30
$a_4a_3^2a_2$	−9	−9	8	5	.	1	5	−14	−2	26	.	−1	−4	18	−16	−36	36
$a_4a_3^2a_1^2$	.	3	2	−4	−6	−20	15	34	−20	−39	18	2	−40	−9	62	39	−72
$a_4a_3a_2^2a_1$	9	6	−8	−19	24	−1	3	16	−36	−51	.	4	−16	9	24	111	−144
$a_4a_3a_2a_1^3$	−3	.	4	7	−12	7	−1	−27	73	68	−12	5	48	−33	−100	−108	240
$a_4a_2^3a_1$	.	.	.	−1	.	.	5	5	−22	−17	.	−5	−8	18	32	17	−72
$a_4a_2^2a_1^3$	.	−2	4	−4	4	.	.	.	−4	−2	−6	12	−12	12	−8	12	−12
$a_4a_2a_1^5$	.	.	−2	8	−18	−1	3	−5	39	16	6	−15	24	−39	10	−76	120
$a_4a_1^8$	.	.	.	.	−2	12	.	−1	5	−54	−20	.	5	−20	45	20	70	−180
$a_3^4a_1$	.	.	.	.	.	−2	.	.	−1	23	8	.	.	4	−21	−14	−18	84
a_3^4	.	.	.	.	.	.	6	.	.	−3	−1	.	.	.	3	2	1	−12
$a_3^3a_2a_1$	3	3	−3	−3	3	−3	6	−3	−3	3	3	−3	−3	6	−3	−3	3
$a_3^2a_2^3$	−3	−3	4	7	−12	7	−18	13	19	−27	−12	13	16	−42	28	37	−48
$a_3^2a_2^2a_1^2$	1	.	−2	−1	6	2	3	−11	−6	18	−2	−5	8	17	−30	−18	40
$a_3^2a_2a_1^4$	.	1	−2	2	−2	−1	3	−5	9	−9	6	−11	8	−3	−6	24	−24
$a_3^2a_1^6$	2	2	1	−4	9	−4	9	−5	−30	54	9	−5	−28	63	−45	−114	180
$a_3^3a_2^2$	9	6	4	1	−6	.	−3	10	13	−45	.	5	−4	−36	70	70	−180
$a_3^3a_1^3$	30	20	15	6	1	.	.	−2	−1	9	.	.	2	3	−17	−9	42
$a_3a_2^3a_1^2$	6	9	4	.	.	1	−3	5	−7	9	−6	10	−4	−6	20	−49	60
$a_3a_2^2a_1^4$	24	24	15	3	.	3	1	−5	14	−30	.	−5	16	−21	.	130	−240
$a_3a_2a_1^6$	81	70	50	17	2	10	5	1	−7	27	.	.	−4	18	−32	−87	252
$a_3^2a_1^6$	252	210	161	70	15	35	21	7	1	−9	.	.	.	−3	16	19	−96
a_2^6	756	630	504	252	72	126	84	36	9	1	.	.	.	.	−2	−1	12
$a_2^4a_1^4$	20	30	15	.	.	6	.	.	.	.	1	−2	2	−2	2	−2	2
$a_2^3a_1^6$	70	81	50	10	.	17	5	.	.	.	2	1	−4	9	−16	25	−36
$a_2^4a_1^4$	228	228	160	52	6	52	24	4	.	.	6	4	1	−6	20	−50	105
$a_2^3a_1^6$	707	660	498	207	42	165	93	27	3	.	20	15	6	1	−8	35	−112
$a_2^2a_1^8$	2128	1932	1526	736	201	532	336	128	26	2	70	56	28	8	1	−10	54
$a_2a_1^{10}$	6300	5670	4620	2460	810	1722	1170	525	145	21	252	210	120	45	10	1	−12
a_1^{12}	18480	16632	13860	7920	2970	5544	3960	1980	660	132	924	792	495	220	66	12	1

Table 3.2

$w=1$. $h_1=a_1$.

$w=2$	a_2	a_1^2
h_2	−1	1
h_1^2		1

Table 3.3

$w=3$	a_3	a_2a_1	a_1^3
h_3	1	−2	1
h_2h_1		−1	1
h_1^3			1

Table 3.4

$w=4$	a_4	a_3a_1	a_2^2	$a_2a_1^2$	a_1^4
h_4	−1	2	1	−3	1
h_3h_1		1	.	−2	1
h_2^2			1	−2	1
$h_2h_1^2$				−1	1
h_1^4					1

Table 3.5

$w=5$	a_5	a_4a_1	a_3a_2	$a_3a_1^2$	$a_2^2a_1$	$a_2a_1^3$	a_1^5
h_5	1	−2	−2	3	3	−4	1
h_4h_1		−1	.	2	1	−3	1
h_3h_2			−1	1	2	−3	1
$h_3h_1^2$				1	.	−2	1
$h_2^2h_1$					1	−2	1
$h_2h_1^3$						−1	1
h_1^5							1

Table 3.6

$w=6$	a_6	a_5a_1	a_4a_2	$a_4a_1^2$	a_3^2	$a_3a_2a_1$	$a_3a_1^3$	a_2^3	$a_2^2a_1^2$	$a_2a_1^4$	a_1^6
h_6	−1	2	2	−3	1	−6	4	−1	6	−5	1
h_5h_1		1	.	−2	.	−2	3	.	3	−4	1
h_4h_2			1	−1	.	−2	2	−1	4	−4	1
$h_4h_1^2$				−1	.	.	2	.	1	−3	1
h_3^2					1	−4	2	.	4	−4	1
$h_3h_2h_1$						−1	1	.	2	−3	1
$h_3h_1^3$							1	.	.	−2	1
h_2^3								−1	3	−3	1
$h_2^2h_1^2$									1	−2	1
$h_2h_1^4$										−1	1
h_1^6											1

Table 3.7

$w=7$	a_7	a_6a_1	a_5a_2	$a_5a_1^2$	a_4a_3	$a_4a_2a_1$	$a_4a_1^3$	$a_3^2a_1$	$a_3a_2^2$	$a_3a_2a_1^2$	$a_3a_1^4$	$a_2^3a_1$	$a_2^2a_1^3$	$a_2a_1^5$	a_1^7
h_7	1	−2	−2	3	−2	6	−4	3	3	−12	5	−4	10	−6	1
h_6h_1		−1	.	2	.	2	−3	1	.	−6	4	−1	6	−5	1
h_5h_2			−1	1	.	2	−2	.	2	−5	3	−3	7	−5	1
$h_5h_1^2$				1	.	.	−2	2	1	−2	3	.	3	−4	1
h_4h_3					−1	2	−1	.	.	−7	3	−2	7	−5	1
$h_4h_2h_1$						1	−1	1	.	−2	2	−1	4	−4	1
$h_4h_1^3$						.	−1	.	.	.	1	.	1	−3	1
$h_3^2h_1$						.	.	2	.	−4	2	.	4	−4	1
$h_3h_2^2$						.	.	.	1	−2	1	−2	5	−4	1
$h_3h_2h_1^2$						.	.	.	.	−1	1	.	2	−3	1
$h_3h_1^4$											1	.	.	−2	1
$h_2^3h_1$												−1	3	−3	1
$h_2^2h_1^3$													1	−2	1
$h_2h_1^5$														−1	1
h_1^7															1

Table 3.8

$w=8$ (i)	a_8	a_7a_1	a_6a_2	$a_5a_1^2$	a_5a_3	$a_4a_3a_1$	$a_4a_1^2$	a_4^2	$a_4a_2a_1$	$a_4a_2^2$	$a_4a_1a_1^2$
h_8	−1	2	2	−3	2	−6	4	1	−6	−3	12
h_7h_1		1	.	−2	.	−2	3	.	−2	.	6
h_6h_2			1	−1	.	−2	2	.	.	−2	5
$h_6h_1^2$				−1	.	.	2	.	.	.	2
h_5h_3					1	−2	1	.	−2	.	4
$h_5h_2h_1$						−1	1	.	.	.	2
$h_5h_1^3$							1	.	.	.	.
h_4^2								1	−4	−2	6
$h_4h_3h_1$									−1	.	2
$h_4h_2^2$										−1	2
$h_4h_2h_1^2$											1

$w=8$ (ii)	$a_6a_1^4$	$a_5^2a_3$	$a_5^2a_1^3$	$a_5a_4^2a_1$	$a_5a_4a_3a_1^2$	$a_5a_3a_1^3$	$a_4^2a_1^4$	$a_4^3a_3^3$	$a_4^3a_1^4$	$a_4^2a_2^2$	$a_4^2a_1^6$
h_8	−5	−3	6	12	−20	6	1	−10	15	−7	1
h_7h_1	−4	.	3	3	−12	5	.	−4	10	−6	1
h_6h_2	−3	−1	1	6	−10	4	1	−7	11	−6	1
$h_6h_1^2$	−3	.	1	.	−6	4	.	−1	6	−5	1
h_5h_3	−2	−2	3	7	−12	4	.	−6	11	−6	1
$h_5h_2h_1$	−2	.	.	2	−5	3	.	−3	7	−5	1
$h_5h_1^3$	−2	.	.	.	−2	3	.	.	3	−4	1
h_4^2	−2	.	4	4	−12	4	1	−6	11	−6	1
$h_4h_3h_1$	−1	.	2	1	−7	3	.	−2	7	−5	1
$h_4h_2^2$	−1	.	.	2	−4	2	1	−5	8	−5	1
$h_4h_2h_1^2$	−1	.	.	.	−2	2	.	−1	4	−4	1
$h_4h_1^4$	−1	.	.	.	.	2	.	.	1	−3	1
$h_3^2h_2$		.	1	4	−6	2	.	−4	8	−5	1
$h_3^2h_1^3$		.	1	.	−4	2	.	.	4	−4	1
$h_3h_2^3$		.	.	.	−2	1	.	−2	5	−4	1
$h_3h_2h_1^3$		.	.	.	−1	1	.	.	2	−3	1
$h_3h_1^5$		.	.	.	.	1	.	.	.	−2	1
h_2^4		.	.	.	.	.	1	−4	6	−4	1
$h_2^3h_1^3$		.	.	.	.	.	.	−1	3	−3	1
$h_2^2h_1^5$		.	.	.	.	.	.	.	1	−2	1
$h_2h_1^7$		.	.	.	.	.	.	.	.	−1	1
h_1^9		.	.	.	.	.	.	.	.	.	1

Table 3.9

$w=9$ (i)	a_9	a_8a_1	a_7a_2	$a_7a_1^2$	a_6a_3	$a_6a_2a_1$	$a_6a_1^3$	a_5a_4	$a_5a_3a_1$	$a_5a_2^2$	$a_5a_2a_1^2$	$a_5a_1^4$	$a_4^2a_1$	$a_4a_3a_2$	$a_4a_2a_1^3$
h_9	1	−2	−2	3	−2	6	−4	−2	6	3	−12	5	3	6	−12
h_8h_1		−1	.	2	.	2	−3	.	2	.	−6	4	1	.	−6
h_7h_2			−1	1	−1	2	−2	.	.	2	−5	3	.	2	−2
$h_7h_1^2$				1	.	.	−2	.	.	.	−2	3	.	.	−2
h_6h_3					.	2	−1	.	2	.	−4	2	.	2	−3
$h_6h_2h_1$					−1	.	−1	.	.	.	−2	2	.	.	.
$h_6h_1^3$						.	−1	.	.	.	.	2	.	.	.
h_5h_4						1	.	−1	2	1	−3	1	2	2	−7
$h_5h_3h_1$						.	.	.	1	.	−2	1	.	.	−2
$h_5h_2^2$						.	.	.	.	.	−2	1	.	.	.
$h_5h_2h_1^2$						.	.	.	.	.	.	1	.	.	.
$h_5h_1^4$						.	.	.	.	.	−1	1	.	.	.
$h_4^2h_1$						.	.	.	.	.	.	1	.	.	.
$h_4h_3h_2$						.	.	.	.	.	.	.	1	.	−4
$h_4h_3h_1^2$						.	.	.	.	.	.	.	.	1	−1
$h_4h_2^2h_1$						.	.	.	.	.	.	.	.	.	−1

$w=9$ (ii)	$a_6a_3^2a_1$	$a_6a_2a_1^3$	$a_6a_1^5$	$a_6^2a_3$	$a_5^2a_3a_1$	$a_5^2a_2^2$	$a_5a_4^2a_1^2$	$a_5a_3^2a_1^2$	$a_5a_2a_1^4$	$a_5a_1^6$	$a_4^2a_1^3$	$a_4^2a_2a_1^2$	$a_4^2a_1^5$	$a_3^3a_2a_1$	$a_3^2a_2^3$	$a_3^2a_2a_1^4$	$a_3^2a_1^6$	$a_2^4a_1$	$a_2^3a_1^3$	$a_2^2a_1^5$	$a_2a_1^7$	a_1^9
h_9	−12	20	−6	1	−12	10	−4	30	−30	7	5	−20	21	−8								1
h_8h_1	−3	12	−5	.	−3	6	.	12	−20	6	1	−10	15	−7								1
h_7h_2	−6	10	−4	.	−3	3	−3	15	−17	5	4	−14	16	−7								1
$h_7h_1^2$	.	6	−4	.	.	3	.	3	−12	5	.	−4	10	−6								1
h_6h_3	−4	8	−3	1	−8	5	−1	18	−19	5	2	−13	16	−7								1
$h_6h_2h_1$	−2	5	−3	.	−1	1	.	6	−10	4	1	−7	11	−6								1
$h_6h_1^3$	.	2	−3	.	.	1	.	.	−6	4	.	−1	6	−5								1
h_5h_4	−5	10	−3	.	−4	6	−2	15	−19	5	3	−13	16	−7								1
$h_5h_3h_1$	.	4	−3	.	−2	3	.	7	−12	4	.	−6	11	−6								1
$h_5h_2^2$	−2	4	−2	.	.	.	−2	7	−8	3	3	−10	12	−6								1
$h_5h_2h_1^2$	.	2	−2	.	.	.	.	2	−5	3	.	−3	7	−5								1
$h_5h_1^4$	.	.	−2	.	.	.	.	.	−2	3	.	.	3	−4								1
$h_4^2h_2$	−2	6	−2	.	.	.	.	4	−12	4	1	−6	11	−6								1
$h_4^2h_1^3$	−2	3	−2	.	−2	4	−1	8	−10	3	2	−9	12	−6								1
$h_4h_3h_2$	−1	2	−1	.	.	2	.	3	−7	2	.	−3	7	−5								1
$h_4h_3h_1^2$	.	1	−1	.	.	2	.	1	−5	2	.	−2	8	−5								1
$h_4h_2^2h_1$	.	.	−1	.	.	1	.	.	−2	2	.	−1	4	−4								1
$h_4h_2h_1^3$	.	.	−1	.	.	1	.	.	−1	2	.	.	1	−3								1
$h_4h_1^5$	.	.	−1	.	.	1	.	.	.	2	.	.	.	−2								1
$h_3^3h_2$	.	.	.	.	−6	3	−1	12	−6	1	.	−8	12	−6								1
$h_3^3h_1^3$	.	.	.	.	−1	1	.	4	−6	1	.	−4	8	−5								1
$h_3^2h_2^2$	.	.	.	.	.	.	.	2	−4	1	.	.	4	−4								1
$h_3^2h_2h_1^2$	.	.	.	.	.	.	−1	3	−3	1	2	−7	9	−5								1
$h_3^2h_1^4$	.	.	.	.	.	.	.	1	−2	1	.	−2	5	−4								1
$h_3h_2^3h_1$	.	.	.	.	.	.	.	.	−1	1	2	.	2	−3								1
$h_3h_2h_1^4$	.	.	.	.	.	.	.	.	.	1	.	−4	3	−3								1
$h_3h_1^6$	.	.	.	.	.	.	.	.	.	.	1	−1	1	−2								1
h_2^5	.	.	.	.	.	.	.	.	.	.	.	.	.	−1								1

Table 3.10

$w=10$ (i)	a_{10}	a_9a_1	a_8a_2	$a_8a_1^2$	a_7a_3	$a_7a_2a_1$	$a_7a_1^3$	a_6a_4	$a_6a_3a_1$	$a_6a_2^3$	$a_6a_2a_1^2$	$a_6a_1^4$	a_5^2	$a_5a_4a_1$
h_{10}	−1	2	2	−3	2	−6	4	2	−6	−3	12	−5	1	−6
h_8h_1		1	.	−2	.	−2	3	.	−2	.	6	−4	.	−2
h_6h_2			1	−1	.	−2	2	.	.	−2	5	−3	.	.
$h_6h_1^2$				−1	.	.	2	.	.	.	2	−3	.	.
h_5h_3					1	−2	1	.	−2	.	4	−2	.	.
$h_5h_2h_1$						−1	1	.	.	.	2	−2	.	.
$h_5h_1^3$							1	.	.	.	.	−2	.	.
h_4h_4								1	−2	−1	3	−1	.	−2
$h_4h_3h_1$									−1	.	2	−1	.	.
$h_4h_2^2$										−1	2	−1	.	.
$h_4h_2h_1^2$											1	−1	.	.
$h_4h_1^4$												−1	.	.
h_3^2													1	−4
$h_3h_2h_1$														−1

$w=10$ (ii)	$a_5a_3a_2$	$a_5a_3a_1^2$	$a_5a_2^2a_1$	$a_5a_2a_1^3$	$a_5a_1^5$	$a_4^2a_2$	$a_4^2a_1^3$	$a_4a_3^2$	$a_4a_3a_2a_1$	$a_4a_2a_1^3$	$a_4a_1^5$	$a_4a_2^3a_1^2$	$a_4a_2a_1^4$	$a_4a_1^6$
h_{10}	−6	12	12	−20	6	−3	6	−3	24	−20	4	−30	30	−7
h_8h_1	.	6	3	−12	5	.	3	.	6	−12	.	−12	20	−6
h_6h_2	−2	2	6	−10	4	−1	1	.	6	−6	3	−15	17	−5
$h_6h_1^2$	.	2	.	−6	4	.	1	.	.	−6	.	−3	12	−5
h_5h_3	−2	3	4	−8	3	.	.	−2	10	−6	.	−12	14	−4
$h_5h_2h_1$	.	.	2	−5	3	.	.	.	2	−2	.	−6	10	−4
$h_5h_1^3$	.	.	.	−2	3	.	.	.	.	−2	.	.	6	−4
h_4h_4	.	4	2	−6	2	−2	3	−1	10	−10	3	−15	16	−4
$h_4h_3h_1$	.	2	.	−4	2	.	.	.	2	−3	.	−4	8	−3
$h_4h_2^2$	.	.	2	−4	2	.	.	.	.	.	2	−7	8	−3
$h_4h_2h_1^2$	.	.	.	−2	2	.	.	.	.	.	.	−2	5	−3
$h_4h_1^4$	.	.	.	.	2	.	.	.	.	.	.	.	2	−3
h_3^2	−4	6	6	−8	2	.	4	.	8	−12	.	−12	16	−4
$h_3h_2h_1$	.	2	1	−3	1	.	2	.	2	−7	.	−5	10	−3
$h_3h_1^3$	−1	1	2	−3	1	.	.	.	2	−2	.	−4	6	−2
$h_2^3h_2$		1	.	−2	1	.	.	.	.	−2	.	.	4	−2
$h_2^3h_1^2$			1	−2	1	.	.	.	.	.	.	−2	4	−2
$h_2^2h_2h_1^2$				−1	1	.	.	.	.	.	.	.	2	−2
$h_2^2h_1^4$					1	.	.	.	.	.	.	.	.	−2
$h_2h_1^6$						−1	1	.	4	−4	2	−8	8	−2
$h_3^2h_1^2$							1	.	.	−4	.	−2	6	−2
$h_3^2h_2^2$								−1	4	−2	.	−4	4	−1
$h_3h_2h_2h_1$									1	−1	.	−2	3	−1
$h_3h_2h_1^3$										−1	.	.	2	−1
$h_3h_1^5$											1	−3	3	−1
$h_2^4h_2$												−1	2	−1
$h_2^4h_1^2$													1	−1
$h_2^3h_1^4$														−1

$w=10$ (iii)	$a_9^2a_1$	$a_9^2a_2^2$	$a_9^2a_1a_1^2$	$a_9^2a_1^4$	$a_9a_2^2a_1$	$a_9a_2^2a_1^3$	$a_9a_2a_1^5$	$a_9a_1^7$	a_8^2	$a_8^2a_1^2$	$a_8a_2^3a_1^4$	$a_8^2a_1^6$	$a_8a_1^8$	a_1^{10}
h_{10}	4	6	−30	15	−20	60	−42	8	−1	15	−35	28	−9	1
h_8h_1	1	.	−12	10	−4	30	−30	7	.	5	−20	21	−8	1
h_6h_2	.	3	−9	6	−12	32	−26	6	−1	11	−25	22	−8	1
$h_6h_1^2$	.	.	−3	6	.	12	−20	6	.	1	−10	15	−7	1
h_5h_3	3	3	−18	8	−10	37	−28	6	.	8	−24	22	−8	1
$h_5h_2h_1$	.	.	−3	3	−3	15	−17	5	.	4	−14	16	−7	1
$h_5h_1^3$	.	.	.	3	.	3	−12	5	.	.	−4	10	−6	1
h_4h_4	2	1	−15	9	−8	34	−28	6	−1	.	−24	22	−8	1
$h_4h_3h_1$	1	.	−8	5	−1	18	−19	5	.	2	−13	16	−7	1
$h_4h_2^2$	.	1	−2	1	−6	16	−14	4	−1	8	−18	17	−7	1
$h_4h_2h_1^2$	.	.	−1	1	.	6	−10	4	.	1	−7	11	−6	1
$h_4h_1^4$	.	.	.	1	.	.	−6	4	.	.	−1	6	−5	1
h_3^2	.	4	−12	9	−12	34	−28	6	.	9	−24	22	−8	1
$h_3h_2h_1$	.	.	−4	6	−2	15	−19	5	.	3	−13	16	−7	1
$h_3h_2h_2$	.	2	−5	3	−7	19	−16	4	.	6	−17	17	−7	1
$h_3h_2h_1^3$	.	.	−2	3	.	7	−12	4	.	.	−6	11	−6	1
$h_3h_2^2h_1$	.	.	.	.	−2	7	−8	3	.	3	−10	12	−6	1
$h_3h_2h_1^3$	.	.	.	.	.	2	−5	3	.	.	−3	7	−5	1
$h_3h_1^5$	.	.	.	.	.	.	−2	3	.	.	.	3	−4	1
$h_2^4h_2$	.	.	−4	4	−4	16	−16	4	−1	7	−17	17	−7	1
$h_2^3h_2^2$	.	.	.	4	.	4	−12	4	.	1	−6	11	−6	1
$h_2^3h_1^4$	2	1	−11	5	−4	22	−18	4	.	4	−16	17	−7	1
$h_2^2h_2h_1h_2$	.	.	−2	2	−1	8	−10	3	.	2	−9	12	−6	1
$h_2^2h_2h_1^3$	.	.	.	2	.	1	−7	3	.	.	−2	7	−5	1
$h_2^2h_1^6$	.	.	.	.	−2	6	−6	2	−1	6	−13	13	−6	1
$h_2h_2^2h_1^2$	.	.	.	.	.	2	−4	2	.	1	−5	8	−5	1
$h_2h_2h_1^5$	.	.	.	.	.	.	−2	2	.	.	−1	4	−4	1
$h_2h_1^8$	.	.	.	.	.	.	.	2	.	.	.	1	−3	1
$h_2^5h_1$	1	.	−6	3	−4	12	−12	3	.	4	−8	12	−6	1
$h_2^5h_2^2$		1	−2	1	−4	10	−8	2	.	4	−12	13	−6	1
$h_2^5h_2h_1^2$			−1	1	.	4	−6	2	.	.	−4	8	−5	1
$h_2^5h_1^4$				1	.	.	−4	2	.	.	.	4	−4	1
$h_2^4h_2^2h_1$					−1	3	−3	1	.	2	−7	9	−5	1
$h_2^4h_2h_1^3$						1	−2	1	.	.	−2	5	−4	1
$h_2^4h_1^5$							−1	1	.	.	.	2	−3	1
$h_2^3h_1^7$								1	.	.	.	.	−2	1
h_2^6									−1	5	−10	10	−5	1
$h_2^6h_1^2$									1	−4	6	−4	1
$h_2^5h_2^2$										−1	3	−3	1	
$h_2^5h_1^4$											1	−2	1	
$h_2^4h_1^6$												−1	1	
$h_2^7h_1^3$													1	

Table 3.11

$w=11$ (i)	a_{11}	$a_{10}a_1$	a_9a_2	$a_9a_1^2$	a_8a_3	$a_8a_2a_1$	$a_8a_1^3$	a_7a_4	$a_7a_3a_1$	$a_7a_2^2$	$a_7a_2a_1^2$	$a_7a_1^4$
h_{11}	1	−2	−2	3	−2	6	−4	−2	6	3	−12	5
$h_{10}h_1$		−1	.	2	.	2	−3	.	2	.	−6	4
h_9h_2			−1	1	.	.	−2	.	.	2	−5	3
$h_9h_1^2$				1	.	.	−2	.	.	.	−2	3
h_8h_3					−1	2	−1	.	2	.	−4	2
$h_8h_2h_1$						1	−1	.	.	.	−2	2
$h_8h_1^3$							−1	.	.	.	.	2
h_7h_4								−1	2	1	−3	1
$h_7h_3h_1$									1	.	−2	1
$h_7h_2^2$										1	−2	1
$h_7h_2h_1^2$											−1	1
$h_7h_1^4$												1

$w=11$ (ii)	a_9a_2	$a_9a_1a_1$	$a_8a_3a_1$	$a_8a_2a_1^2$	$a_8a_2^2a_1$	$a_8a_2a_1^3$	$a_8a_1^5$	$a_7^2a_1$	$a_7a_4a_2$	$a_7a_4a_1^2$	$a_7a_3^2$	$a_7a_3a_2a_1$	$a_7a_2a_1^3$
h_{11}	−2	6	6	−12	−12	20	−6	3	6	−12	3	−24	20
$h_{10}h_1$	.	2	.	−6	−3	12	−5	1	.	−6	.	−6	12
h_9h_2	.	.	2	−2	−6	10	−4	.	2	−2	.	−6	6
$h_9h_1^2$	.	.	.	−2	.	6	−4	.	.	−2	.	.	6
h_8h_3	.	.	2	−3	−4	8	−3	.	.	.	2	−10	6
$h_8h_2h_1$	.	.	.	.	−2	5	−3	.	.	.	.	−2	2
$h_8h_1^3$	.	.	.	.	.	2	−3	.	.	.	.	.	2
h_7h_4	.	2	.	−4	−2	6	−2	.	.	−3	.	−4	6
$h_7h_3h_1$	.	.	.	−2	.	4	−2	.	2	.	.	−2	3
$h_7h_2^2$	.	.	.	.	−2	4	−2	.	.	.	.	.	.
$h_7h_2h_1^2$	.	.	.	.	.	2	−2	.	.	.	.	.	.
$h_7h_1^4$	.	.	.	.	.	.	−2	.	.	.	.	.	.
h_6h_5	−1	2	.	−3	−3	4	−1	2	2	−7	1	−10	10
$h_6h_4h_1$		1	2	−2	−1	3	−1	.	.	−2	.	.	4
$h_6h_3h_2$			.	−1	−2	3	−1	.	.	.	.	−2	2
$h_6h_3h_1^2$			1	−1	.	3	−1	.	.	.	.	.	2
$h_6h_2^2h_1$					−1	2	−1	.	.	.	.	.	.
$h_6h_2h_1^3$						2	−1	.	.	.	.	.	.
$h_6h_1^5$						1	−1	.	.	.	.	.	.
$h_5^2h_1$								1	.	−4	.	−4	6
$h_5h_4h_3$									1	−1	.	−2	2
$h_5h_4h_2h_1$										−1	.	.	2
$h_5h_4h_1^3$											1	−4	2
$h_5h_3^2h_2$												−1	1
$h_5h_3^2h_1^2$													1

$w=11$ (iii)	$a_9a_2^2$	$a_9a_2^2a_1^2$	$a_9a_2a_1^4$	$a_9a_1^6$	$a_8^2a_3$	$a_8^2a_2a_1$	$a_8^2a_1^3$	$a_8a_3^2a_1$	$a_8a_2^3$	$a_8a_2^2a_1^3$	$a_8a_2a_1^5$	$a_8a_1^7$
h_{11}	−4	30	−30	7	3	−12	10	−12	−12	60	−30	20	−60	42	−8
$h_{10}h_1$	.	12	−20	6	.	−3	6	−3	.	24	−20	4	−30	30	−7
h_9h_2	−3	15	−17	5	.	−3	3	.	−6	18	−12	12	−32	26	−6
$h_9h_1^3$	.	3	−12	5	.	.	3	.	.	6	−12	.	−12	20	−6
h_8h_3	.	12	−14	4	1	−2	1	−6	−3	24	−11	6	−27	22	−5
$h_8h_2h_1$	.	6	−10	4	.	−1	1	.	.	6	−6	3	−15	17	−5
$h_8h_1^3$	.	.	−6	4	.	.	1	.	.	.	−6	.	−3	12	−3
h_7h_4	−2	9	−11	3	2	−6	4	−7	−5	30	−15	10	−32	24	−5
$h_7h_3h_1$	.	4	−8	3	.	.	.	−2	.	10	−6	.	−12	14	−4
$h_7h_2^2$	−2	7	−8	3	.	.	.	.	−2	4	−2	6	−16	14	−4
$h_7h_2h_1^2$	.	2	−5	3	.	.	.	.	.	2	−2	.	−6	10	−4
$h_7h_1^4$	.	.	−2	3	.	.	.	.	.	.	−2	.	.	6	−4
h_6h_5	−1	12	−13	3	.	−4	6	−2	−4	24	−17	8	−29	24	−5
$h_6h_4h_1$	.	2	−6	2	.	−2	3	−1	.	10	−10	3	−15	16	−4
$h_6h_3h_2$	.	4	−6	2	.	.	.	.	−2	5	−3	4	−12	11	−3
$h_6h_3h_1^2$	.	.	−4	2	.	.	.	.	.	2	−3	.	−4	8	−3
$h_6h_2^2h_1$	.	2	−4	2	.	.	.	.	.	.	.	2	−7	8	−3
$h_6h_2h_1^3$	.	.	−2	2	.	.	.	.	.	.	.	.	−2	5	−3
$h_6h_1^5$	.	.	.	2	.	.	.	.	.	.	.	.	.	2	−3
$h_5^2h_1$	.	6	−8	2	.	.	4	.	.	8	−12	.	−12	16	−4
$h_5h_4h_3$	−1	4	−4	1	.	−2	2	.	−2	9	−7	5	−15	13	−3
$h_5h_4h_2h_1$	.	1	−3	1	.	.	2	.	.	2	−7	.	−8	10	−3
$h_5h_4h_1^3$	.	4	−4	1	.	.	.	−2	.	8	−4	.	−8	8	−2
$h_5h_3^2h_2$	.	2	−3	1	.	.	.	.	.	2	−2	.	−4	6	−2
$h_5h_3^2h_1^2$	.	.	−2	1	.	.	.	.	.	.	−2	.	.	4	−2
$h_5h_2^2h_1^2$	.	3	−3	1	.	.	.	.	.	.	.	2	−6	6	−2
$h_5h_2h_1^4$	.	1	−2	1	.	.	.	.	.	.	.	.	−2	4	−2
$h_5h_1^6$	.	.	−1	1	.	.	.	.	.	.	.	.	.	2	−2
$h_4^2h_3$	.	.	.	1	.	−2	1	−4	−2	14	−6	4	−14	10	−2
$h_4^2h_2h_1$						−1	1	.	.	4	−4	2	−8	8	−2
$h_4^2h_1^3$							1	.	.	4	−4	.	−2	6	−2
$h_4h_3^2h_1$								−1	.	4	−2	.	−4	4	−1
$h_4h_3h_2^2$									−1	2	−1	2	−5	4	−1
$h_4h_3h_2h_1^2$										1	−1	.	−2	3	−1
$h_4h_2^3h_1$											−1	1	−3	3	−1
$h_4h_2^2h_1^3$												.	−1	2	−1
$h_4h_2h_1^5$														1	−1
$h_4h_1^7$															−1

Table 3.11 (*cont.*)

$w=11$ (iv)	$a_2^3a_8$	$a_3^3a_2^2$	$a_3^2a_5^2a_1$	$a_3^2a_2a_1^3$	$a_3^2a_1^5$	$a_3a_2^4$	$a_3a_2^2a_1^3$	$a_3a_2^3a_1^2$	$a_3a_2a_1^5$	$a_3a_1^7$	$a_2^5a_1$	$a_2^4a_1^3$	$a_2^3a_1^5$	$a_2^2a_1^7$	$a_2a_1^9$	a_1^{11}
h_{11}	−4	10	30	−60	21	5	−60	105	−56	9	−6	35	−56	36	−10	1
$h_{10}h_1$	.	4	6	−30	15	.	−20	60	−42	8	−1	15	−35	28	−9	1
h_9h_2	−1	1	12	−22	10	4	−34	60	−37	7	−5	25	−41	29	−9	1
$h_9h_1^2$	.	1	.	−12	10	.	−4	30	−30	7	.	5	−20	21	−8	1
h_8h_3	−3	6	18	−35	12	1	−34	67	−39	7	−2	21	−40	29	−9	1
$h_8h_2h_1$	.	.	3	−9	6	.	−12	32	−25	6	−1	11	−25	22	−8	1
$h_8h_1^3$	.	.	.	−3	6	.	.	12	−20	6	.	1	−10	15	−7	1
h_7h_4	.	6	9	−33	13	3	−29	64	−39	7	−4	22	−40	29	−9	1
$h_7h_3h_1$	.	3	3	−18	8	.	−10	37	−28	6	.	8	−24	22	−8	1
$h_7h_2^2$	.	.	3	−6	3	3	−18	32	−22	5	−4	18	−30	23	−8	1
$h_7h_2h_1^2$	.	.	.	−3	3	.	−3	15	−17	5	.	4	−14	16	−7	1
$h_7h_1^4$	.	.	.	.	3	.	.	3	−12	5	.	.	−4	10	−6	1
h_6h_5	−2	3	15	−30	13	2	−33	64	−39	7	−3	22	−40	29	−9	1
$h_6h_4h_1$	.	2	1	−15	9	.	−8	34	−28	6	−1	9	−24	22	−8	1
$h_6h_3h_2$	−1	1	8	−13	5	1	−19	37	−24	5	−2	15	−29	23	−8	1
$h_6h_3h_1^2$	.	1	.	−8	5	.	−1	18	−19	5	.	2	−13	16	−7	1
$h_6h_2^2h_1$	.	.	1	−2	1	.	−6	16	−14	4	−1	8	−18	17	−7	1
$h_6h_2h_1^3$	.	.	.	−1	1	.	.	6	−10	4	.	1	−7	11	−6	1
$h_6h_1^5$	.	.	.	.	1	.	.	.	−6	4	.	.	−1	6	−5	1
$h_5^2h_1$	.	.	4	−12	9	.	−12	34	−28	6	.	9	−24	22	−8	1
$h_5h_4h_2$	.	.	4	−10	6	2	−17	34	−24	5	−3	16	−29	23	−8	1
$h_5h_4h_1^2$	.	.	.	−4	6	.	−2	15	−19	5	.	3	−13	16	−7	1
$h_5h_3^2$	−2	3	11	−20	7	.	−20	42	−26	5	.	12	−28	23	−8	1
$h_5h_3h_2h_1$	.	.	2	−5	3	.	−7	19	−16	4	.	6	−17	17	−7	1
$h_5h_3h_1^3$	.	.	.	−2	3	.	.	7	−12	4	.	.	−6	11	−6	1
$h_5h_2^3$	.	.	.	.	.	2	−9	15	−11	3	−3	13	−22	18	−7	1
$h_5h_2^2h_1^2$	.	.	.	.	.	.	−2	7	−8	3	.	3	−10	12	−6	1
$h_5h_2h_1^4$	.	.	.	.	.	.	.	2	−5	3	.	.	−3	7	−5	1
$h_5h_1^6$	.	.	.	.	.	.	.	.	−2	3	.	.	.	3	−4	1
$h_4^2h_3$	.	4	4	−20	8	1	−14	39	−26	5	−2	13	−28	23	−8	1
$h_4^2h_2h_1$	.	.	.	−4	4	.	−4	16	−16	4	−1	7	−17	17	−7	1
$h_4^2h_1^3$	.	.	.	.	4	.	.	4	−12	4	.	1	−6	11	−6	1
$h_4h_3^2h_1$	.	2	1	−11	5	.	−4	22	−18	4	.	4	−16	17	−7	1
$h_4h_3h_2^2$	.	.	2	−4	2	1	−9	18	−13	3	−2	11	−21	18	−7	1
$h_4h_3h_2h_1^2$	.	.	.	−2	2	.	−1	8	−10	3	.	2	−9	12	−6	1
$h_4h_3h_1^4$	.	.	.	.	2	.	.	1	−7	3	.	.	−2	7	−5	1
$h_4h_2^4$	.	.	.	.	.	.	−2	6	−6	2	−1	6	−13	13	−6	1
$h_4h_2^3h_1^2$	.	.	.	.	.	.	.	2	−4	2	.	1	−5	8	−5	1
$h_4h_2h_1^4$	.	.	.	.	.	.	.	.	−2	2	.	.	−1	4	−4	1
$h_4h_1^8$	.	.	.	.	.	.	.	.	.	2	.	.	.	1	−3	1
$h_3^3h_2$	−1	1	6	−9	3	.	−12	24	−15	3	.	8	−20	18	−7	1
$h_3^3h_1^2$		1	.	−6	3	.	.	12	−12	3	.	.	−8	12	−6	1
$h_3^2h_2^2h_1$			1	−2	1	.	−4	10	−8	2	.	4	−12	13	−6	1
$h_3^2h_2h_1^3$				−1	1	.	.	4	−6	2	.	.	−4	8	−5	1
$h_3^2h_1^5$					1	.	.	.	−4	2	.	.	.	4	−4	1
$h_3h_2^4$						1	−4	6	−4	1	−2	9	−16	14	−6	1
$h_3h_2^3h_1^2$							−1	3	−3	1	.	2	−7	9	−5	1
$h_3h_2^2h_1^4$								1	−2	1	.	.	−2	5	−4	1
$h_3h_2h_1^6$									−1	1	.	.	.	2	−3	1
$h_3h_1^8$										1	.	.	.	.	−2	1
$h_2^6h_1$											−1	5	−10	10	−5	1
$h_2^5h_1^3$												1	−4	6	−4	1
$h_2^4h_1^5$													−1	3	−3	1
$h_2^3h_1^7$														1	−2	1
$h_2^2h_1^9$															−1	1
h_1^{11}																1

Table 3.12

$w=12$ (i)	a_{12}	$a_{11}a_1$	$a_{10}a_2$	$a_{10}a_1^2$	a_9a_3	$a_9a_2a_1$	$a_9a_1^3$	a_8a_4	$a_8a_3a_1$	$a_8a_2^2$	$a_8a_2a_1^2$	$a_8a_1^4$	a_7a_5	$a_7a_4a_1$	$a_7a_3a_2$
h_{12}	−1	2	2	−3	2	−6	4	2	−6	−3	12	−5	2	−6	−6
$h_{11}h_1$		1	.	−2	.	−2	3	.	−2	.	6	−4	.	−2	.
$h_{10}h_2$			1	−1	.	−2	2	.	.	−2	5	−3	.	.	−2
$h_{10}h_1^2$				−1	.	.	2	.	.	.	2	−3	.	.	.
h_9h_3					1	−2	1	.	−2	.	4	−2	.	.	−2
$h_9h_2h_1$						−1	1	.	.	.	2	−2	.	.	.
$h_9h_1^3$							1	.	.	.	.	−2	.	.	.
h_8h_4								1	−2	.	3	−1	.	−2	.
$h_8h_3h_1$								.	−1	.	2	−1	.	.	.
$h_8h_2^2$								.	.	−1	2	−1	.	.	.
$h_8h_2h_1^2$															
$h_8h_1^4$											1	−1	.	.	.
h_7h_5												−1	.	.	.
$h_7h_4h_1$													1	−2	.
$h_7h_3h_2$														−1	.
$h_7h_3h_1^2$															−1

$w=12$ (ii)	$a_7a_3a_1^2$	$a_7a_2^2a_1$	$a_7a_2a_1^3$	$a_7a_1^5$	a_6^2	$a_6a_4a_2$	$a_6a_4a_1^2$	$a_6a_3^2$	$a_6a_3a_2a_1$	$a_6a_3a_1^3$	$a_6a_2^3$	$a_6a_2^2a_1^2$	$a_6a_2a_1^4$	$a_6a_1^6$
h_{12}	12	12	−20	6	1	−6	−6	12	−3	24	−20	4	−30	30	−7
$h_{11}h_1$	6	3	−12	5	.	−2	.	6	.	6	−12	.	−12	20	−6
$h_{10}h_2$	2	0	−10	4	.	.	−2	2	.	6	−6	3	−15	17	−5
$h_{10}h_1^2$	2	.	−6	4	.	.	.	2	.	.	−6	.	−3	12	−5
h_9h_3	3	4	−8	3	.	.	.	.	−2	10	−6	.	−12	14	−4
$h_9h_2h_1$	.	2	−5	3	.	.	.	.	.	2	−2	.	−6	10	−4
$h_9h_1^3$	.	.	−2	3	.	.	.	.	.	.	−2	.	.	6	−4
h_8h_4	4	2	−6	2	.	.	−2	3	.	4	−6	2	−9	11	−3
$h_8h_3h_1$	2	.	−4	2	.	.	.	.	.	2	−3	.	−4	8	−3
$h_8h_2^2$	.	2	−4	2	.	.	.	.	.	.	.	2	−7	8	−3
$h_8h_2h_1^2$			−2	2	.	.	.	.	.	.	.	.	−2	5	−3
$h_8h_1^4$	.	.	.	2	.	.	.	.	.	.	.	.	.	2	−3
h_7h_5	3	3	−4	1	.	−2	.	4	.	4	−6	.	−6	8	−2
$h_7h_4h_1$	2	1	−3	1	.	.	.	2	.	.	−4	.	−2	6	−2
$h_7h_3h_2$	1	2	−3	1	.	.	.	.	.	2	−2	.	−4	6	−2
$h_7h_3h_1^2$	1	.	−2	1	.	.	.	.	.	.	−2	.	.	4	−2
$h_7h_2^2h_1$		1	−2	1	.	.	.	.	.	.	.	.	−2	4	−2
$h_7h_2h_1^3$			−1	1	.	.	.	.	.	.	.	.	.	2	−2
$h_7h_1^5$				1	.	.	.	.	.	.	.	.	.	.	−2
h_6^2					1	−4	−4	6	−2	12	−8	2	−12	10	2
$h_6h_5h_1$						−1	.	2	.	2	−3	.	−3	4	−1
$h_6h_4h_2$							−1	1	.	2	−2	1	−4	4	−1
$h_6h_4h_1^2$								1	.	.	−2	.	−1	4	−1
$h_6h_3^2$								.	−1	4	−2	.	−4	3	−1
$h_6h_3h_2h_1$										1	−2	.	−4	4	−1
$h_6h_3h_1^3$											−1	.	−2	2	−1
$h_6h_2^3$												1	.	3	−1
$h_6h_2^2h_1^2$													−3	3	−1
$h_6h_2h_1^4$													−1	2	−1
$h_6h_1^6$														1	−1

$w=12$ (iii)	$a_6^2a_2$	$a_6^2a_1^2$	$a_6a_4a_2$	$a_6a_4a_2a_1$	$a_6a_4a_1^3$	$a_6a_3^2a_1$	$a_6a_3a_2^2$	$a_6a_3a_2a_1^2$	$a_6a_3a_1^4$	$a_6a_2^3a_1$	$a_6a_2^2a_1^3$	$a_6a_2a_1^5$	$a_6a_1^7$	a_4^3	$a_4^2a_3a_1$
h_{12}	−3	6	−6	24	−20	12	12	−60	30	−20	60	−42	8	−1	12
$h_{11}h_1$	.	3	.	6	−12	3	.	−24	20	−4	30	−30	7	.	3
$h_{10}h_2$	−1	1	.	6	−6	.	6	−18	12	−12	32	−26	6	.	.
$h_{10}h_1^2$	.	1	.	.	−6	.	.	−6	12	.	12	−20	6	.	.
h_9h_3	.	.	−2	4	−2	6	3	−24	11	−6	27	−22	5	.	3
$h_9h_2h_1$	.	.	.	2	−2	.	.	−6	6	−3	15	−17	5	.	.
$h_9h_1^3$	.	.	.	.	−2	.	.	.	6	.	3	−12	5	.	.
h_8h_4	.	.	−2	6	−4	4	2	−18	10	−6	22	−18	4	−1	8
$h_8h_3h_1$	.	.	.	.	.	2	.	−10	6	.	12	−14	4	.	1
$h_8h_2^2$	.	.	.	.	.	.	2	−4	2	−6	16	−14	4	.	.
$h_8h_2h_1^2$	.	.	.	.	.	.	.	−2	2	.	6	−10	4	.	.
$h_8h_1^4$	.	.	.	.	.	.	.	.	2	.	.	−6	4	.	.
h_7h_5	−2	3	−2	10	−10	3	7	−24	14	−10	27	−20	4	.	4
$h_7h_4h_1$	.	.	.	2	−3	.	.	−4	6	−2	9	−11	3	.	2
$h_7h_3h_2$	.	.	.	.	.	2	.	−5	3	−4	12	−11	3	.	.
$h_7h_3h_1^2$	.	.	.	.	.	.	.	−2	3	.	4	−8	3	.	.
$h_7h_2^2h_1$	.	.	.	.	.	.	.	.	.	−2	7	−8	3	.	.
$h_7h_2h_1^3$	.	.	.	.	.	.	.	.	.	.	2	−5	3	.	.
$h_7h_1^5$	.	.	.	.	.	.	.	.	.	.	.	−2	3	.	.
h_6^2	.	4	.	8	−12	4	.	−24	16	−4	24	−20	4	.	.
$h_6h_5h_1$	.	2	.	2	−7	1	.	−10	10	−1	12	−13	3	.	.
$h_6h_4h_2$	.	.	.	2	−2	.	.	−4	4	−2	8	−8	2	.	.
$h_6h_4h_1^2$	.	.	.	.	−2	.	.	−8	4	.	2	−6	2	.	.
$h_6h_3^2$	.	.	.	.	.	2	.	−8	4	.	8	−8	2	.	.
$h_6h_3h_2h_1$	.	.	.	.	.	.	.	−2	2	.	4	−6	2	.	.
$h_6h_3h_1^3$	.	.	.	.	.	.	.	.	.	−2	6	−6	2	.	.
$h_6h_2^3$	.	.	.	.	.	.	.	.	.	.	2	−4	2	.	.
$h_6h_2^2h_1^2$	.	.	.	.	.	.	.	.	.	.	.	−2	2	.	.
$h_6h_2h_1^4$	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.
$h_6h_1^6$	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$h_5^2h_2$	−1	1	.	4	−4	.	4	−10	6	−6	14	−10	2	.	.
$h_5^2h_1^2$	.	1	.	.	−4	.	.	−4	6	.	6	−8	2	.	.
$h_5h_4h_3$	.	.	−1	2	−1	2	1	−7	3	−2	7	−5	1	.	2
$h_5h_4h_2h_1$	.	.	.	1	−1	.	.	−2	2	−1	4	−4	1	.	.
$h_5h_4h_1^3$	.	.	.	.	−1	.	.	.	2	.	1	−3	1	.	.
$h_5h_3^2h_1$	.	.	.	.	.	.	1	−4	2	.	4	−4	1	.	.
$h_5h_3h_2^2$	.	.	.	.	.	.	.	−2	1	−2	5	−4	1	.	.
$h_5h_3h_2h_1^2$	.	.	.	.	.	.	.	−1	1	.	3	−3	1	.	.
$h_5h_3h_1^4$	.	.	.	.	.	.	.	.	.	−1	3	−3	1	.	.
$h_5h_2^3h_1$	.	.	.	.	.	.	.	.	.	.	1	−2	1	.	.
$h_5h_2^2h_1^3$	.	.	.	.	.	.	.	.	.	.	.	−1	1	.	.
$h_5h_2h_1^5$	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.
$h_5h_1^7$	.	.	.	.	.	.	.	.	.	.	.	.	.	−1	6
$h_4^2h_3h_1$	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1

Table 3.12 (*cont.*)

$w=12$ (iv)	$a_1^3a_2^3$	$a_1^3a_2a_1^3$	$a_3^3a_1^4$	$a_4a_2^3a_1$	$a_4a_2^3a_1^3$	$a_4a_2a_3^2a_1$	$a_4a_2a_3a_1^3$	$a_4a_3a_1^5$	$a_4a_5^4$	$a_4a_3^3a_1^3$	$a_4a_3^3a_1^4$	$a_4a_4a_1^5$	$a_5a_1^5$	a_6^4	$a_3^3a_3a_1$	$a_3^3a_1^3$
h_{12}	6	−30	15	12	−30	−60	120	−42	−5	60	−105	56	−9	1	−20	20
$h_{11}h_1$	.	−12	10	.	−12	−12	60	−30	.	20	−60	42	−8	.	−4	10
$h_{10}h_2$	3	−9	6	3	−3	−24	44	−20	−4	34	−60	37	−7	.	−4	4
$h_{10}h_1^2$	.	−3	6	.	−3	.	24	−20	.	4	−30	30	−7	.	.	4
h_9h_3	.	−6	3	6	−12	−24	50	−18	.	24	−52	32	−6	1	−14	11
$h_9h_2h_1$	.	−3	3	.	.	−6	18	−12	.	12	−32	26	−6	.	−1	1
$h_9h_1^3$	.	.	3	.	.	.	6	−12	.	.	−12	20	−6	.	.	1
h_8h_4	4	−15	6	3	−18	−24	62	−22	−4	31	−59	34	−6	.	−6	12
$h_8h_3h_1$	.	−2	1	.	−6	−3	24	−11	.	6	−27	22	−5	.	−3	6
$h_8h_2^2$	1	−2	1	.	.	−6	12	−6	−3	18	−32	22	−5	.	.	.
$h_8h_2h_1^2$	.	−1	1	.	.	.	6	−6	.	3	−15	17	−5	.	.	.
$h_8h_1^4$	.	.	1	.	.	.	.	−6	.	.	−3	12	−5	.	.	.
h_7h_5	.	−12	8	4	−12	−24	58	−24	.	26	−56	34	−6	.	−6	9
$h_7h_4h_1$	.	−6	4	.	−7	−5	30	−15	.	10	−32	24	−5	.	.	6
$h_7h_3h_2$	.	.	.	2	−2	−10	16	−6	.	12	−26	18	−4	.	−3	3
$h_7h_3h_1^2$	.	.	.	.	−2	.	10	−6	.	.	−12	14	−4	.	.	3
$h_7h_2^2h_1$	.	.	.	.	.	−2	4	−2	.	6	−16	14	−4	.	−3	.
$h_7h_2h_1^3$	.	.	.	.	.	.	2	−2	.	.	−6	10	−4	.	.	.
$h_7h_1^5$	.	.	.	.	.	.	.	−2	.	.	.	6	−4	.	.	.
h_6^2	4	−12	9	4	−6	−24	52	−24	−4	30	−56	34	−6	1	−12	8
$h_6h_5h_1$	.	−4	6	.	−2	−4	24	−17	.	8	−29	24	−5	.	−2	3
$h_6h_4h_2$	2	−5	3	1	−1	−10	20	−10	−3	18	−31	20	−4	.	−2	2
$h_6h_4h_1^2$	.	−2	3	.	−1	.	10	−10	.	3	−15	16	−4	.	.	2
$h_6h_3^2$	.	.	.	2	−3	−8	16	−6	.	8	−20	14	−3	1	−10	6
$h_6h_3h_2h_1$	.	.	.	.	.	−2	5	−3	.	4	−12	11	−3	.	−1	1
$h_6h_3h_1^3$	.	.	.	.	.	.	2	−3	.	.	−4	8	−3	.	.	1
$h_6h_2^3$	.	.	.	.	.	.	.	.	−2	9	−15	11	−3	.	.	.
$h_6h_2^2h_1^2$	.	.	.	.	.	.	.	.	.	2	−7	8	−3	.	.	.
$h_6h_2h_1^4$	.	.	.	.	.	.	.	.	.	.	−2	5	−3	.	.	.
$h_6h_1^6$	.	.	.	.	.	.	.	.	.	.	.	2	−3	.	.	.
$h_5^2h_2$	.	−4	4	.	.	−8	20	−12	.	12	−28	20	−4	.	.	.
$h_5^2h_1^2$	.	.	4	.	.	.	8	−12	.	.	−12	16	−4	.	.	6
$h_5h_4h_3$	.	−4	2	2	−7	−9	26	−10	.	10	−25	16	−3	.	−4	.
$h_5h_4h_2h_1$	.	−2	2	.	.	−2	9	−7	.	5	−15	13	−3	.	.	.
$h_5h_4h_1^3$	.	.	2	.	.	.	2	−7	.	.	−5	10	−3	.	.	.
$h_5h_3^2h_1$	.	.	.	.	−2	.	8	−4	.	.	−8	8	−2	.	−2	3
$h_5h_3h_2^2$	.	.	.	.	.	−2	4	−2	.	4	−10	8	−2	.	.	.
$h_5h_3h_2h_1^2$	.	.	.	.	.	.	2	−2	.	.	−4	6	−2	.	.	.
$h_5h_3h_1^4$	.	.	.	.	.	.	.	−2	.	.	.	4	−2	.	.	.
$h_5h_2^4h_1$	.	.	.	.	.	.	.	.	.	2	−6	6	−2	.	.	.
$h_5h_2^3h_1^3$	.	.	.	.	.	.	.	.	.	.	−2	4	−2	.	.	.
$h_5h_2^2h_1^4$	.	.	.	.	.	.	.	.	.	.	.	2	−2	.	.	.
$h_5h_2h_1^6$	.	.	.	.	.	.	.	.	.	.	.	.	−2	.	.	.
$h_5h_1^7$	.	.	.	.	.	.	.	.	.	.	.	.	−2	.	.	.
h_4^3	3	−9	3	.	−12	−12	36	−12	−3	18	−33	18	−3	.	.	8
$h_4^2h_3h_1$	.	−2	1	.	−4	−2	14	−6	.	4	−14	10	−2	.	.	4
$h_4^2h_2^2$	1	−2	1	.	.	−4	8	−4	−2	10	−16	10	−2	.	.	.
$h_4^2h_2h_1^2$	.	−1	1	.	.	.	4	−4	.	2	−8	8	−2	.	.	.
$h_4^2h_1^4$	.	.	1	.	.	.	.	−4	.	.	−2	6	−2	.	.	.
$h_4h_3^2h_2$	.	.	.	1	−1	−4	6	−2	.	4	−8	5	−1	.	−2	2
$h_4h_3^2h_1^2$	.	.	.	.	−1	.	4	−2	.	.	−4	4	−1	.	.	2
$h_4h_3h_2^3$	.	.	.	.	.	−1	2	−1	.	2	−5	4	−1	.	.	.
$h_4h_3h_2^2h_1^3$	.	.	.	.	.	.	1	−1	.	.	−2	3	−1	.	.	.
$h_4h_3h_2h_1^5$	.	.	.	.	.	.	.	−1	.	.	.	2	−1	.	.	.
$h_4h_3^3$	.	.	.	.	.	.	.	.	−1	4	−6	4	−1	.	.	.
$h_4h_3^2h_2^3$	.	.	.	.	.	.	.	.	.	1	−3	3	−1	.	.	.
$h_4h_3^2h_2^2h_1^4$	.	.	.	.	.	.	.	.	.	.	−1	2	−1	.	.	.
$h_4h_3^2h_2h_1^6$	.	.	.	.	.	.	.	.	.	.	.	1	−1	.	.	.
$h_4h_3^2h_1^7$	.	.	.	.	.	.	.	.	.	.	.	.	−1	.	.	.
h_4^4	.	.	.	.	.	.	.	.	.	.	.	.	.	1	−8	4
$h_4^3h_3h_2h_1$	.	.	.	.	.	.	.	.	.	.	.	.	.	.	−1	1
$h_4^3h_2^3h_1^3$	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1

Table 3.12 (*cont.*)

$w=12$ (v)	$a_5^3a_2^3$	$a_5^3a_2^2a_1^2$	$a_5^3a_3a_2a_1^4$	$a_5^3a_1^6$	$a_5a_4^2a_1$	$a_5a_3^2a_1^3$	$a_5a_3^3a_1^3$	$a_5a_3a_1^7$	$a_5a_1^8$	a_2^6	$a_5^5a_1^3$	$a_5^4a_1^4$	$a_5^3a_1^6$	$a_5^2a_2^2$	$a_5a_1^{10}$	a_1^{11}
h_{12}	−10	90	−105	28	30	−140	168	−72	10	1	−21	70	−84	45	−11	1
$h_{11}h_1$	.	30	−60	21	5	−60	105	−56	9	.	−6	35	−56	36	−10	1
$h_{10}h_2$	−6	36	−45	15	20	−80	102	−50	8	1	−16	50	−63	37	−10	1
$h_{10}h_1^2$	.	6	−30	15	.	−20	60	−42	8	.	−1	15	−35	28	−9	1
h_9h_3	−4	54	−62	17	13	−84	111	−52	8	.	−10	45	−62	37	−10	1
$h_9h_2h_1$	.	12	−22	10	4	−34	60	−37	7	.	−5	25	−41	29	−9	1
$h_9h_1^3$	.	.	−12	10	.	−4	30	−30	7	.	.	5	−20	21	−8	1
h_8h_4	−3	39	−61	18	14	−76	108	−52	8	1	−13	46	−62	37	−10	1
$h_8h_3h_1$	.	18	−35	12	1	−34	67	−39	7	.	−2	21	−40	29	−9	1
$h_8h_2^2$	−3	12	−15	6	12	−44	58	−32	6	1	−12	36	−47	30	−9	1
$h_8h_2h_1^2$	.	3	−9	6	.	−12	32	−26	6	.	−1	11	−25	22	−8	1
$h_8h_1^4$	.	.	−3	6	.	.	12	−20	6	.	.	1	−10	15	−7	1
h_7h_5	−6	42	−58	18	17	−80	108	−52	8	.	−12	46	−62	37	−10	1
$h_7h_4h_1$	.	9	−33	13	3	−29	64	−39	7	.	−4	22	−40	29	−9	1
$h_7h_3h_2$	−3	21	−26	8	10	−47	65	−34	6	.	−8	32	−46	30	−9	1
$h_7h_3h_1^2$	.	3	−18	8	.	−10	37	−28	6	.	.	8	−24	22	−8	1
$h_7h_2^2h_1$	.	3	−6	3	3	−18	32	−22	5	.	−4	18	−30	23	−8	1
$h_7h_2h_1^3$	.	.	−3	3	.	−3	15	−17	5	.	.	4	−14	16	−7	1
$h_7h_1^5$	.	.	.	3	.	.	3	−12	5	.	.	.	−4	10	−6	1
h_6^2	−2	48	−58	18	12	−80	108	−52	8	1	−12	46	−62	37	−10	1
$h_6h_5h_1$	.	15	−30	13	2	−33	64	−39	7	.	−3	22	−40	29	−9	1
$h_6h_4h_2$	−1	16	−24	9	8	−42	62	−34	6	1	−10	33	−46	30	−9	1
$h_6h_4h_1^2$	.	1	−15	9	.	−8	34	−28	6	.	−1	9	−24	22	−8	1
$h_6h_3^2$	−1	34	−37	10	5	−50	72	−36	6	.	−4	28	−45	30	−9	1
$h_6h_3h_2h_1$	.	8	−13	5	1	−19	37	−24	5	.	−2	15	−29	23	−8	1
$h_6h_3h_1^3$	.	.	−8	5	.	−1	18	−19	5	.	.	2	−13	16	−7	1
$h_6h_2^3$	−1	3	−3	1	6	−22	30	−18	4	1	−9	26	−35	24	−8	1
$h_6h_2^2h_1^2$	.	1	−2	1	.	−6	16	−14	4	.	−1	8	−18	17	−7	1
$h_6h_2h_1^4$	.	.	−1	1	.	.	6	−10	4	.	.	1	−7	11	−6	1
$h_6h_1^6$	.	.	.	1	.	.	.	−6	4	.	.	.	−1	6	−5	1
$h_5^2h_2$	−4	16	−21	9	12	−46	62	−34	6	.	−9	33	−46	30	−9	1
$h_5^2h_1^2$	.	4	−12	9	.	−12	34	−28	6	.	.	9	−24	22	−8	1
$h_5h_4h_3$	−2	23	−35	11	7	−45	69	−36	6	.	−6	29	−45	30	−9	1
$h_5h_4h_2h_1$	.	4	−10	6	2	−17	34	−24	5	.	−3	16	−29	23	−8	1
$h_5h_4h_1^3$	.	.	−4	6	.	−2	15	−19	5	.	.	3	−13	16	−7	1
$h_5h_3^2h_1$	.	11	−20	7	.	−20	42	−26	5	.	.	12	−28	23	−8	1
$h_5h_3h_2^2$	−2	7	−8	3	7	−26	35	−20	4	.	−6	23	−34	24	−8	1
$h_5h_3h_2h_1^2$	.	2	−5	3	.	−7	19	−16	4	.	.	6	−17	17	−7	1
$h_5h_3h_1^4$	.	.	−2	3	.	.	7	−12	4	.	.	.	−6	11	−6	1
$h_5h_2^3h_1$	.	.	.	.	2	−9	15	−11	3	.	−3	13	−22	18	−7	1
$h_5h_2^2h_1^3$	.	.	.	.	.	−2	7	−8	3	.	.	3	−10	12	−6	1
$h_5h_2h_1^5$	.	.	.	.	.	.	2	−5	3	.	.	.	−3	7	−5	1
$h_5h_1^7$	.	.	.	.	.	.	.	−2	3	.	.	.	.	3	−4	1
h_4^3	.	12	−36	12	6	−36	66	−36	6	1	−9	30	−45	30	−9	1
$h_4^2h_3h_1$	.	4	−20	8	1	−14	39	−26	5	1	−2	13	−28	23	−8	1
$h_4^2h_2^2$	.	4	−8	4	4	−20	32	−20	4	1	−8	24	−34	24	−8	1
$h_4^2h_2h_1^2$	.	.	−4	4	.	−4	16	−16	4	.	−1	7	−17	17	−7	1
$h_4^2h_1^4$	.	.	.	4	.	.	4	−12	4	.	.	1	−6	11	−6	1
$h_4h_3^2h_2$	−1	12	−16	5	4	−26	40	−22	4	.	−4	20	−33	24	−8	1
$h_4h_3^2h_1^2$	.	1	−11	5	.	−4	22	−18	4	.	.	4	−16	17	−7	1
$h_4h_3h_2^2h_1$	.	2	−4	2	1	−9	18	−13	3	.	−2	11	−21	18	−7	1
$h_4h_3h_2h_1^3$	.	.	−2	2	.	−1	8	−10	3	.	.	2	−9	12	−6	1
$h_4h_3h_1^5$	.	.	.	2	.	.	1	−7	3	.	.	.	−2	7	−5	1
$h_4h_2^4$	.	.	.	.	2	−8	12	−8	2	1	−7	10	−26	19	−7	1
$h_4h_2^3h_1^2$	.	.	.	.	.	−2	6	−6	2	.	−1	6	−13	13	−6	1
$h_4h_2^2h_1^4$	.	.	.	.	.	.	2	−4	2	.	.	1	−5	8	−5	1
$h_4h_2h_1^6$	.	.	.	.	.	.	.	−2	2	.	.	.	−1	4	−4	1
$h_4h_1^8$	.	.	.	.	.	.	.	.	2	.	.	.	.	1	−3	1
h_3^4	.	24	−24	6	.	−32	48	−24	4	.	.	16	−32	24	−8	1
$h_3^3h_2h_1$	.	6	−9	3	.	−12	24	−15	3	.	.	8	−20	18	−7	1
$h_3^3h_2^2$	.	.	−6	3	.	.	12	−12	3	.	.	.	−8	12	−6	1
$h_3^3h_1^3$	−1	3	−3	1	.	−4	10	−8	2	.	.	4	−12	13	−6	1
$h_3^2h_2^3h_1$	.	1	−2	1	.	−4	10	−8	2	.	.	4	−12	13	−6	1
$h_3^2h_2h_1^3$	.	.	−1	1	.	.	4	−6	2	.	.	.	−4	8	−5	1
$h_3^2h_1^6$	.	.	.	1	.	.	.	−4	2	.	.	.	.	4	−4	1
$h_3h_2^5h_1$	.	.	.	.	1	−4	6	−4	1	.	−2	9	−16	14	−6	1
$h_3h_2^4h_1^2$	.	.	.	.	.	−1	3	−3	1	.	.	2	−7	9	−5	1
$h_3h_2^3h_1^4$	.	.	.	.	.	.	1	−2	1	.	.	.	−2	5	−4	1
$h_3h_2^2h_1^6$	.	.	.	.	.	.	.	−1	1	.	.	.	.	2	−3	1
$h_3h_2h_1^8$	.	.	.	.	.	.	.	.	1	.	.	.	.	.	−2	1
h_2^6										1	−6	15	−20	15	−6	1
$h_2^5h_1^2$											−1	5	−10	10	−5	1
$h_2^4h_1^4$												1	−4	6	−4	1
$h_2^3h_1^6$													−1	3	−3	1
$h_2^2h_1^8$														1	−2	1
$h_2h_1^{10}$															−1	1
h_1^{12}																1

MISCELLANEA

Moment constants for the distribution of range in normal samples

By H. O. HARTLEY and E. S. PEARSON

Given a random sample of n observations $x_1, \ldots, x_n$ arranged in ascending order of magnitude, drawn from a normal population with unit variance, we denote the range by $w = x_n - x_1$ and its probability density distribution by $f_n(w)$.

Tables of the probability integral of range

$$P_n(W) = \int_0^W f_n(w)\,dw,$$

as well as of its percentage points, are available (see, for example, Pearson & Hartley, 1942). However, the numerous applications of this statistic frequently lead to distributions derived from that of range. For example, the mean of k independent ranges is often used for the estimation of the parental standard deviation σ. The distributions of such derived statistics could be obtained from the table of $P_n(W)$ by numerical processes of quadrature (see, for example, Lord, 1947). In order to avoid the labour involved in such procedures it is often possible to use analytic approximations to $f_n(w)$ (Cox, 1949; Florin, 1950; Patnaik, 1950) which permit the analytic evaluation of the distribution of the derived statistics. In order to facilitate the choice of such approximations and to provide checks on their accuracy it was felt that a systematic tabulation of the moments of $f_n(w)$ would be helpful, apart from the independent value of putting them on record.†

n	μ_1'	σ	μ_2	β_1	β_2	μ_3	μ_4	μ_5	μ_6	κ_4	κ_5	κ_6
2	1·12838	0·8525	0·72676	0·9906	3·869	0·6166	2·044	4·46	13·3	0·459	−0·03	−1·22
3	1·69257	0·8884	0·78922	0·4176	3·286	0·4531	2·047	3·44	11·1	0·178	−0·13	−0·41
4	2·05875	0·8798	0·77407	0·2735	3·188	0·3561	1·910	2·70	9·4	0·113	−0·06	−0·15
5	2·32593	0·8641	0·74661	0·2174	3·169	0·3008	1·766	2·22	8·1	0·094	−0·02	−0·07
6	2·53441	0·8480	0·71916	0·1892	3·168	0·2653	1·638	1·91	7·2	0·087	+0·00	−0·05
7	2·70436	0·8332	0·69424	0·1742	3·174	0·2415	1·530	1·69	6·4	0·084	+0·02	−0·03
8	2·84720	0·8198	0·67213	0·1657	3·184	0·2243	1·438	1·53	5·9	0·083	+0·02	−0·02
9	2·97003	0·8078	0·65262	0·1608	3·191	0·2114	1·359	1·41	5·4	0·081	+0·03	−0·02
10	3·07751	0·7971	0·63531	0·1580	3·200	0·2013	1·292	1·31	5·0	0·081	+0·03	−0·00
11	3·17287	0·7873	0·61984	0·1564	3·205	0·1930	1·231	1·23	4·7	0·079	+0·03	−0·01
12	3·25846	0·7785	0·60601	0·1560	3·213	0·1863	1·180	1·16	4·4	0·078	+0·04	−0·00
13	3·33598	0·7704	0·59353	0·1559	3·220	0·1806	1·134					
14	3·40676	0·7630	0·58217	0·1561	3·225	0·1755	1·093					
15	3·47183	0·7562	0·57186	0·1568	3·231	0·1712	1·057					
16	3·53198	0·7499	0·56237	0·1576	3·237	0·1674	1·024					
17	3·58788	0·7441	0·55363	0·1588	3·242	0·1641	0·994					
18	3·64006	0·7386	0·54554	0·1598	3·248	0·1611	0·967					
19	3·68896	0·7335	0·53802	0·1612	3·254	0·1585	0·942					
20	3·73495	0·7287	0·53097	0·1627	3·259	0·1561	0·919					

Definitions

$$\mu_1' = \int_0^\infty w f_n(w)\,dw;\ \mu_r = \int_0^\infty (w-\mu_1')^r f_n(w)\,dw;$$

$$\sigma = \sqrt{\mu_2};\ \beta_1 = \mu_3^2/\mu_2^3,\ \beta_2 = \mu_4/\mu_2^2;$$

$$\kappa_4 = \mu_4 - 3\mu_2^2,\ \kappa_5 = \mu_5 - 10\mu_3\mu_2,$$

$$\kappa_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3.$$

The table above was derived from the 5-decimal manuscript of the Pearson-Hartley table by numerical quadrature. Its accuracy and scope is limited by the fact that the higher moments require an increasing accuracy in the tails of $f_n(w)$. From this point of view it was found most convenient to introduce suitably chosen 'working means' w_n^*, depending on n, and to evaluate the moments μ_r^* about them by applying quadrature to the formula

$$\mu_r^* = (-w_n^*)^r + r\int_0^\infty (w-w_n^*)^{r-1}(1-P_n(w))\,dw.$$

† Certain results up to $n = 12$ have already been tabulated by Grubbs & Weaver (1947). We note certain discrepancies between their results and the present table. The method of computation is not given in their paper.

From these moments, the moment-constants defined in the table were computed. The values of μ_1' are exact; the remaining values may be in error in the last figure quoted.

For expert help in the computational work we are indebted to Mr T. Vickers of the Mathematics Division of the National Physical Laboratories, while the moments about the mean and the cumulants were calculated from the μ^* by Miss E. R. Fitch.

REFERENCES

Cox, D. R. (1949). *J.R. Statist. Soc.* B, **11**, 101.
Florin, H. (1950). *Communications of the Royal Flemish Academy* (Science Series), **12**, 6.
Grubbs, F. E. & Weaver, C. L. (1947). *J. Amer. Statist. Ass.* **42**, 224.
Lord, E. (1947). *Biometrika*, **34**, 41.
Patnaik, P. B. (1950). *Biometrika*, **37**, 78.
Pearson, E. S. & Hartley, H. O. (1942). *Biometrika*, **32**, 300.

Rapid methods for estimating correlation coefficients

By L. N. CHOWN and P. A. P. MORAN
Balliol College and the Institute of Statistics, Oxford University

If $\hat{\theta}_1$ is a consistent estimator of a population parameter θ and $\hat{\theta}_2$ is a most efficient estimator of θ, the efficiency of $\hat{\theta}_1$ is defined to be $\mathrm{var}(\hat{\theta}_2)/\mathrm{var}(\hat{\theta}_1)$. In practice, however, the times required to calculate $\hat{\theta}_1$ and $\hat{\theta}_2$ may be very different, and from the point of view of the cost of computation, especially if an indefinitely large amount of data is available, it may be better to assess the efficiency by multiplying the above quantity by t_2/t_1, where t_1 is the time or labour required to calculate $\hat{\theta}_1$, and t_2 the time or labour required to calculate $\hat{\theta}_2$. In the present paper we consider two cases of this kind. In the first part we describe what appears to be the most rapid method of estimating a product-moment correlation coefficient yet invented, and in the second part we consider a method of estimating the first-order serial correlation coefficient in a sample of a time series generated by a simple Markov scheme.

Suppose we have a sample $\{x_i, y_i\}\,(i = 1, \ldots, n)$ from a bivariate normal distribution with unknown standard deviations and correlation coefficient ρ. The sample product-moment correlation coefficient r is an asymptotically most efficient estimator of ρ but requires a fair amount of computation. Kendall's rank-correlation coefficient t is easier to calculate for small n, and $\frac{2}{\pi}\sin^{-1}t$ is an estimator of ρ which is consistent and has an efficiency $9\pi^{-2}$ in large samples for values of ρ near zero. However, for large n, it is probably more complicated to calculate than r. It is, however, possible to find another correlation coefficient, of ranking type, which is extremely easy to calculate and has an efficiency, when used as an estimator of ρ, for ρ near zero, of about 0·33.

Write $\mathrm{sgn}\,x = \pm 1$ as $x \gtrless 0$ and $\mathrm{sgn}\,0 = 0$, and define

$$g = \frac{1}{n-1}\sum_{i=1}^{n-1} a_{i,i+1}b_{i,i+1}, \tag{1}$$

where $$a_{ij} = \mathrm{sgn}\,(x_i - x_j), \quad b_{ij} = \mathrm{sgn}\,(y_i - y_j).$$

g depends on the order in which the pairs (x_i, y_i) are given, but if this is random,

$$E(g) = E(a_{i,i+1}b_{i,i+1}).$$

But $x_i - x_{i+1}$ and $y_i - y_{i+1}$ are jointly distributed in a normal distribution with correlation coefficient ρ and so (compare Kendall, 1948, p. 114)

$$E(g) = E(a_{i,i+1}b_{i,i+1}) = \frac{2}{\pi}\sin^{-1}\rho. \tag{2}$$

As an estimator of ρ we could therefore take $r_g = \sin\frac{1}{2}\pi g$. From (1) we also have

$$(n-1)^2E(g^2) = E\left\{\sum_{i=1}^{n-1} a_{i,i+1}^2b_{i,i+1}^2 + 2\sum_{i=1}^{n-2} a_{i,i+1}b_{i,i+1}a_{i+1,i+2}b_{i+1,i+2} + \sum\sum_{|i-j|>1} a_{i,i+1}b_{i,i+1}a_{j,j+1}b_{j,j+1}\right\}. \tag{3}$$

Clearly $E(a_{i,i+1}^2) = E(b_{i,i+1}^2) = 1$, and in his investigation of the variance of the rank-correlation coefficient t, Kendall shows (1948, p. 115) that

$$E(a_{i,i+1}b_{i,i+1}a_{i+1,i+2}b_{i+1,i+2}) = \left(\frac{2}{\pi}\sin^{-1}\rho\right)^2 - \left(\frac{2}{\pi}\sin^{-1}\tfrac{1}{2}\rho\right)^2 + \frac{1}{9}.$$

If $|i-j|>1$ we also have

$$E(a_{i,i+1}b_{i,i+1}a_{j,j+1}b_{j,j+1}) = [E(a_{i,i+1}b_{i,i+1})]^2 = \left(\frac{2}{\pi}\sin^{-1}\rho\right)^2.$$

Substituting in (3) we have, after a little reduction and the subtraction of $\{E(g)\}^2$,

$$\text{var}\,(g) = \frac{1}{n-1}\left[1-\left(\frac{2}{\pi}\sin^{-1}\rho\right)^2\right] + \frac{2(n-2)}{(n-1)^2}\left[\frac{1}{9}-\left(\frac{2}{\pi}\sin^{-1}\tfrac{1}{2}\rho\right)^2\right]. \qquad (4)$$

Thus if $\rho = 0$,

$$\text{var}\,(g) = \frac{11n-13}{9(n-1)^2} \doteqdot \frac{11}{9(n-1)}.$$

Since $|\sin^{-1}\tfrac{1}{2}\rho| > |\tfrac{1}{2}\sin^{-1}\rho|$, we have

$$\frac{1}{9}-\left(\frac{2}{\pi}\sin^{-1}\tfrac{1}{2}\rho\right)^2 \leqslant \frac{1}{9}\left\{1-\left(\frac{2}{\pi}\sin^{-1}\rho\right)^2\right\} \doteqdot \tfrac{1}{9}(1-g^2).$$

Hence we have the approximate inequality

$$\text{var}\,(g) \leqslant \frac{11n-13}{9(n-1)^2}(1-g^2).$$

In the neighbourhood of $\rho = 0$ we have $\delta r_g = \tfrac{1}{2}\pi\,\delta g$ and so

$$\text{var}\,(r_g) = \tfrac{1}{4}\pi^2\,\text{var}\,g \doteqdot \frac{11\pi^2}{36(n-1)}.$$

Then since r and r_g are consistent estimators of ρ, we see that the efficiency of r_g as an estimator of ρ, is, in the neighbourhood of $\rho = 0$,

$$\frac{\text{var}\,(r)}{\text{var}\,(r_g)} = \frac{36}{11\pi^2} = 0{\cdot}33,$$

and thus the use of r_g involves throwing away about two-thirds of the information. Nevertheless, if the pairs of observations are in random order, the labour of calculating r_g from $3n$ pairs of observations is very much less than that of calculating r from n pairs.

Cochran (1937) describes another rapid method of estimating ρ.* If $\bar{x}$ and $\bar{y}$ are the means of the x's and y's he finds the proportion of cases in which $\text{sgn}\,(x_i-\bar{x})\,\text{sgn}\,(y_i-\bar{y}) > 0$, and equates this to

$$\tfrac{1}{2} + \pi^{-1}\sin^{-1}\rho.$$

He shows that in the neighbourhood of $\rho = 0$, the efficiency of this method, when used as an estimator of ρ, is 0·405. However, the present method appears to be much faster and therefore preferable in cases where the pairs of values are given in random order.

Now suppose that one of the variates, y say, is given in a prescribed order. We can still test whether the x's are independent by showing that in the universe of $n!$ equiprobable permutations of the x's, g has expectation zero and a variance depending on the y's but with a usable upper bound. For in such a case we obviously have

$$E(a_{i,i+1}) = 0,$$

$$E(a_{i,i+1}a_{i+1,i+2}) = -\tfrac{1}{3},$$

and so

$$E(g) = 0,$$

$$(n-1)^2\,\text{var}\,(g) = (n-1) - \frac{2}{3}\sum_{i=1}^{n-2} b_{i,i+1}b_{i+1,i+2}.$$

The sum on the right-hand side could be calculated in any given case. However, we shall always have

$$\left|\sum_{i=1}^{n-2} b_{i,i+1}b_{i+1,i+2}\right| \leqslant (n-2),$$

* [This method has sometimes been described as 'Sheppard's method' of estimating a correlation coefficient; it follows from the work in his 1899 paper, *Phil. Trans.* A, **192**, 101. Ed.]

and so $$(n-1)^2 \operatorname{var}(g) \leqslant (n-1) + \tfrac{8}{3}(n-2) = \tfrac{1}{3}(5n-7),$$

and thus $$\operatorname{var}(g) \leqslant \frac{5n-7}{3(n-1)^2}, \tag{5}$$

which will often provide a reasonable test.

The exact distribution of g for small n and all permutations of x equiprobable has not been found either for a fixed or a random order of the y's. Such a distribution would be interesting but probably not of much practical value, for if n is small the loss of information is so large that the present method is scarcely worth while, and it would be better to use a product-moment or a ranking correlation coefficient.

A very similar type of method may be used for the estimation of serial correlations. Consider a stationary time series generated by a relationship of the form

$$(x_{i+1} - m) = \rho(x_i - m) + \epsilon_i, \tag{6}$$

where m is unknown and $\{\epsilon_i\}$ is a sequence of independent random variables which are normally distributed with zero means and the same, but unknown, standard deviations. Write

$$X_i = \operatorname{sgn}(x_{i+1} - x_i) \operatorname{sgn}(x_{i+2} - x_{i+1}) \quad (i = 1, \ldots, n-2).$$

Put $$S = \frac{1}{n-2} \sum_{i=1}^{n-2} X_i. \tag{7}$$

Then if $\rho = 0$, we have $$E(X_i) = -\tfrac{1}{3}, \tag{8}$$

and $$E(S^2) = (n-2)^{-2} E\left(\sum_1^{n-2} X_i\right)^2$$

$$= (n-2)^{-2} E\left\{\sum_1^{n-2} X_i^2 + 2\sum_1^{n-3} X_i X_{i+1} + 2\sum_1^{n-4} X_i X_{i+2} + \sum\sum_{|i-j|>2} X_i X_j\right\}.$$

Now $E(X_i^2) = 1$, and if $|i-j| > 2$, $\quad E(X_i X_j) = \{E(X_i)\}^2 = \frac{1}{9}$.

By enumerating all of the possible 4! and 5! permutations respectively it may be shown that $E(X_i X_{i+1}) = 0$ and $E(X_i X_{i+2}) = \frac{2}{15}$. Then

$$E(S^2) = (n-2)^{-2}\{(n-2) + \tfrac{4}{15}(n-4) + \tfrac{1}{9}(n-4)(n-5)\}$$

$$= \frac{5n^2 + 12n - 38}{45(n-2)^2},$$

and $$\operatorname{var}(S) = \frac{32n-58}{45(n-2)^2} \doteq \frac{32}{45n}. \tag{9}$$

The calculation of S is simple and quick, and if n is reasonably large provides us with a useful rough test of the serial independence of the sequence $(x_1, \ldots, x_n)$. It may be noticed that this method is related to the calculation of Kendall (1946, p. 382) of the expected number of peaks in a stationary time series, the difference being that we here use both peaks and 'troughs', and also consider the variance.

We now consider what happens when S is used to provide an estimator of ρ in a process generated by a relation of type (6), with $\rho \neq 0$. Consider any sequence (x_1, x_2, x_3). Then $(x_2 - x_1)$ and $(x_3 - x_2)$ are easily seen to be jointly normally distributed with a correlation coefficient $-\frac{1}{2}(1-\rho)$. It follows that

$$\text{pr.}\{x_2 - x_1 > 0,\ x_3 - x_2 > 0\} = \text{pr.}\{x_2 - x_1 < 0,\ x_3 - x_2 < 0\}$$

$$= \frac{1}{2}\left\{1 - \frac{1}{\pi}\cos^{-1}\left(\frac{\rho-1}{2}\right)\right\}.$$

Thus $$E(S) = E(X_i) = 1 - \frac{2}{\pi}\cos^{-1}\left(\frac{\rho-1}{2}\right). \tag{10}$$

We therefore take as our estimator of ρ, $\hat{\rho} = 1 + 2\sin\frac{1}{2}\pi S$.

We have not been able to find the variance of S when $\rho \neq 0$, but as it is clearly a continuous function of ρ in the neighbourhood of $\rho = 0$ we can find the efficiency of $\hat{\rho}$ as an estimator of ρ relative to a maximum-likelihood estimator when ρ is small. $\hat{\rho}$ is clearly a consistent estimator, and when ρ is near zero we can write

$$\delta\hat{\rho} = \pi \cos\tfrac{1}{2}\pi S \, \delta S,$$

so that $$\text{var}(\hat{\rho}) \doteqdot \pi^2 \cos^2 \tfrac{1}{6}\pi \,\text{var}(S) \doteqdot \frac{8\pi^2}{15n}.$$

On the other hand, the maximum-likelihood estimator of ρ is also consistent and has, near $\rho = 0$, a variance equal to $\frac{1}{n} + O\left(\frac{1}{n^2}\right)$. Consequently, the relative efficiency of $\hat{\rho}$ is

$$\frac{15}{8\pi^2} = 0{\cdot}19.$$

This method of estimating ρ is not only rapid but gives us an unbiased estimate of a function of ρ, and is independent of m and the variance of ϵ_t. It may therefore be useful if we have to estimate ρ from a large number of short series. To do this we add together the values of S, each of which is an unbiased estimator of the right-hand side of equation (10). We are indebted to Mr E. J. Williams of the Australian C.S.I.R.O. for drawing our attention to the need for a test of this kind. Although we have not found the variance of S when $\rho \neq 0$, if we are estimating ρ from a number of short series we can find an estimate of var (S) from the observed variance between the values obtained.

One of us (L. N. C.) is indebted to the Department of Scientific and Industrial Research for a maintenance grant.

REFERENCES

COCHRAN, W. G. (1937). *J.R. Statist. Soc.* **100**, 69.
KENDALL, M. G. (1946). *Advanced Theory of Statistics*, **2**. London: Charles Griffin and Co.
KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Charles Griffin and Co,

We are indebted to Mr D. E. Barton of University College, London for the following notes:

The moments of S as defined in equation (7) may be derived for second and higher orders without recourse to the heavy labour involved in enumerating large numbers of permutations. For these moments are seen to be functions of terms such as $E(X_{i_1} X_{i_2} \dots X_{i_n})$ of a simple form, and these expectations may be evaluated in the following manner.

We take $i_1 < i_2 < \dots < i_n$ and find that:

(*a*) If $|i_r - i_{r+1}| > 2$ for $r = 1, \dots, n-1$, then

$$E(X_{i_1} X_{i_2} \dots X_{i_n}) = \{E(X_i)\}^n.$$

(*b*) If $(i_1, i_2, \dots, i_n)$ contains a group $(i_r, i_{r+1}, \dots, i_{r+t})$ with the following properties,

(i) $i_{r+s+1} - i_{r+s} = 1$ for all $S = 0, 1, \dots, t-1$,

(ii) neither $i_{r+t+1} - i_{r+t} = 1$ nor $i_r - i_{r-1} = 1$,

(iii) either $i_{r+t+1} - i_{r+t} > 2$ or $i_r - i_{r-1} > 2$,

then $$E(X_{i_1} X_{i_2} \dots X_{i_n}) = 0.$$

(*c*) If neither (*a*) nor (*b*) are true, then $E(X_{i_1} X_{i_2} \dots X_{i_n})$ is a product of such terms as

$$P_{m+1} = E\{\text{sgn}(x_{i+1} - x_i)\,\text{sgn}(x_{i+2} - x_{i+1}) \dots \text{sgn}(x_{i+m} - x_{i+m-1})\}.$$

Further, $P_m = a_{m+1,1}/m!$, where $a_{m+1,1}$ is shown simply to be determined by the recurrence relations

$$a_{m+1,r} = \sum_{t=r}^{m} a_{m,t} - \sum_{t=1}^{r-1} a_{m,t} \quad (r = 1, 2, \dots, m+1),$$

where $a_{1,1} = 1$.

Thus $$P_3 = -\tfrac{1}{3}, \quad P_5 = \tfrac{2}{15}, \quad P_7 = -\tfrac{17}{315}, \quad P_9 = \tfrac{62}{2835},$$

and, from the symmetry of the recurrence, it is seen that $P_m = 0$ for even values of m.

Mean and variance of an entry in a contingency table

By W. L. STEVENS

Faculdade de Ciências Econômicas e Administrativas, Universidade de São Paulo

Patnaik (1948), in a paper which examines the power function of the test for independence in a 2×2 table, derives certain formulae for the mean and variance of a tabular entry which may, under other circumstances, prove to be misleading.

Using Patnaik's notation, we may represent the contingency table by

a	c	m
b	d	n
r	s	N

If m and n are arbitrarily chosen, if the probabilities are equal that an event classified in either the first or second row will fall in the first column, and if we consider only those results for which r and s have given values, then it can be shown that the mean and variance of a are

$$\mu = mr/N, \tag{1}$$

$$\sigma^2 = mnrs/\{N^2(N-1)\}. \tag{2}$$

When the table is not 'independent', let us designate these probabilities and their complements respectively by

$$p_1, \quad q_1 = 1-p_1,$$

$$p_2, \quad q_2 = 1-p_2,$$

and write

$$\psi = p_1 q_2/(p_2 q_1) \neq 1. \tag{3}$$

Then it can be shown that the probabilities in the distribution of a are *proportional* to

$$\frac{m!n!r!s!}{N!a!b!c!d!}\psi^a. \tag{4}$$

Patnaik seeks the mean and variance of this distribution. Following him, we replace the first factor of (4) by the ordinate of the normal distribution with mean and variance* given by (1) and (2) and write the second factor as $\exp(a \log \psi)$. We deduce that the distribution of a is given by

$$\frac{1}{\sigma\sqrt{(2\pi)}} \exp\left\{-\frac{(a-\mu-\sigma^2\log\psi)^2}{2\sigma^2}\right\}, \tag{5}$$

whence we conclude that

$$\text{mean } (a) = \mu + \sigma^2 \log \psi, \tag{6}$$

$$\text{variance } (a) = \sigma^2. \tag{7}$$

It is difficult to detect any flaw in this argument, but that the result is erroneous can readily be shown. Let us suppose, for example, that $p_1 = 0{\cdot}02$ and $p_2 = 0{\cdot}98$. Suppose further that $m = n = r = s = \frac{1}{2}N$. It will be noted that we are not endeavouring to obtain a peculiar result by imposing unnatural restrictions on r and s. On the contrary, the values chosen are the most probable.

Now suppose that N and hence m tend to infinity. It is evident that, notwithstanding the restriction on r and s, the proportion a/m will tend, in the probability sense, to $p_1 = 0{\cdot}02$. Hence a/N tends to $0{\cdot}01$ and the formula for the mean should therefore give

$$\text{mean } (a) = 0{\cdot}01N. \tag{8}$$

Using Patnaik's formula, we find, when N is large,

$$\left.\begin{aligned} \mu &= 0{\cdot}25N, \qquad \sigma^2 = 0{\cdot}0625N, \\ \psi &= (\tfrac{1}{49})^2, \qquad \log\psi = -7{\cdot}7836 \\ \text{mean } (a) &= \textit{minus } 0{\cdot}2365N. \end{aligned}\right\} \tag{9}$$

Apart from being wrong, the result is evidently absurd—the mean of positive values cannot be negative.

* It is open to question whether there is any advantage in using $N^2(N-1)$ as the denominator of expression (2) instead of the simpler N^3.

We must therefore re-examine the argument by which these formulae were obtained. It is observed that, keeping the marginal proportions fixed and allowing N to tend to ∞, the variance given in (2) also tends to ∞. Hence, unless $\psi = 1$, the term $\sigma^2 \log \psi$ in equation (5) will tend to $\pm\infty$. Consequently, in order to obtain finite values for $(a - \mu - \sigma^2 \log \psi)$, we must suppose that $|a - \mu| \to \infty$. This implies that the pertinent values of

$$\frac{m!n!r!s!}{N!a!b!c!d!} \tag{10}$$

are those in which the values of a are increasingly remote from the mean. Now, although it is legitimate to replace the *bulk* of the distribution (10) by a normal distribution with the same mean and variance, it is quite wrong to suppose that this approximation will be valid in an increasingly remote portion of a tail of the distribution.

It appears then that, for Patnaik's derivation to be valid, we must suppose, not only that $N \to \infty$, but also that $\psi \to 1$. Since he is studying the power function of the test for independence, i.e. the test of the hypothesis that $\psi = 1$, the use of the formulae, *in this particular context*, is entirely valid. In other words, (6) and (7) are valid approximations in the neighbourhood of $\psi = 1$. As, however, this limitation does not appear in his derivation, the formulae may yet prove a trap for someone who uses them for other values of ψ.

We give below, without proof*, the formulae for the mean and variance of a. Let four quantities, π_1, π_2, π_3 and π_4, be defined by the relations

$$\left.\begin{aligned} \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1, \\ \pi_1 + \pi_2 &= r/N, \\ \pi_1 + \pi_3 &= m/N, \\ \pi_1\pi_4/\pi_2\pi_3 &= \psi. \end{aligned}\right\} \tag{11}$$

Then, in the limit,

$$\left.\begin{aligned} \text{mean } (a) &= N\pi_1, \\ \text{variance } (a) &= \frac{N}{\frac{1}{\pi_1} + \frac{1}{\pi_2} + \frac{1}{\pi_3} + \frac{1}{\pi_4}}. \end{aligned}\right\} \tag{12}$$

The variance is seen to be a natural extension of the variance of a binomial, when the latter is written in the form

$$\frac{N}{\frac{1}{\pi} + \frac{1}{1-\pi}}.$$

To return to the example examined earlier, we have $\pi_1 = \pi_4 = 0{\cdot}01$ and $\pi_2 = \pi_3 = 0{\cdot}49$. Hence

$$\text{mean } (a) = 0{\cdot}01N,$$

as already given in equation (8), and

$$\text{variance } (a) = 49N/9802 = 0{\cdot}005N \quad \text{approximately.}$$

It appears that the variance is about one-half of what it would be, if no restrictions were placed on the marginal subtotals.

To check Patnaik's approximations, let us consider a less extreme example: $\pi_1 = \pi_4 = 0{\cdot}2$, $\pi_2 = \pi_3 = 0{\cdot}3$ and all marginal subtotals equal, as before. Then, N being large, we have by the two methods:

	Proposed	Patnaik
Mean (a)	$0{\cdot}2N$	$0{\cdot}19932N$
Variance (a)	$0{\cdot}06N$	$0{\cdot}0625N$

It appears that the results are in fair agreement in this instance, even although $\psi(= \frac{4}{9})$ is 'a good way' from unity.

* See Kendall (1947), Vol. 1, p. 312, for a similar result.

In conclusion, we suggest that the problem of approximating to mean, variance and higher cumulants of an entry in a 2×2 table with $\psi \neq 1$, is one which merits further study; the 'well-worn' topic of the 2×2 table, far from being exhausted, is proving itself to be inexhaustible.

REFERENCES

KENDALL, M. G. (1947). *The Advanced Theory of Statistics*, **1**. London: Charles Griffin and Co. Ltd.

PATNAIK, P. B. (1948). The power function of the test for the difference between two proportions in a 2×2 table. *Biometrika*, **35**, 157–73.

The Wishart distribution derived by solving simultaneous linear differential equations

BY H. S. STEYN, *University of Pretoria*

Let us first consider the multivariate normal probability function

$$\phi \equiv \phi(x_1, x_2, \ldots, x_k) = C \exp\{-\tfrac{1}{2}(a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + a_{kk}x_k^2)\}, \tag{1}$$

where the function inside the bracket is a positive definite quadratic form. On differentiation we have

$$-\frac{\partial}{\partial x_i}\phi = (a_{1i}x_1 + a_{2i}x_2 + \ldots + a_{ik}x_k)\,\phi \quad (i = 1, 2, \ldots, k). \tag{2}$$

Now, if $M \equiv M(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the moment-generating function of (1) it is obvious from

$$M = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} e^{\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k}\,\phi\, dx_1 dx_2 \ldots dx_k,$$

that $\alpha_i M$ is the m.g.f. of $-\dfrac{\partial}{\partial x_i}\phi$,

and $\dfrac{\partial}{\partial \alpha_i} M$ is the m.g.f. of $x_i\phi$. (3)

Thus, it follows from (2) that

$$\alpha_i M = a_{1i}\frac{\partial}{\partial \alpha_1}M + a_{2i}\frac{\partial}{\partial \alpha_2}M + \ldots + a_{ik}\frac{\partial}{\partial \alpha_k}M \quad (i = 1, 2, \ldots, k).$$

Solving the k equations in $\dfrac{1}{M}\dfrac{\partial}{\partial \alpha_r}M\,(r = 1, 2, \ldots, k)$, we have

$$\frac{1}{M}\frac{\partial}{\partial \alpha_i}M = \frac{1}{|A|}\{|A_{1i}|\alpha_1 + |A_{2i}|\alpha_2 + \ldots + |A_{ii}|\alpha_i + \ldots + |A_{ik}|\alpha_k\},$$

where the determinant

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{12} & a_{22} & \ldots & a_{2k} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ a_{1k} & a_{2k} & \ldots & a_{kk} \end{vmatrix}$$

and $|A_{ij}| = |A_{ji}|$ is the co-factor of a_{ij} or a_{ji}. Integrating we find

$$M = \exp\left\{\frac{1}{2}\frac{1}{|A|}(|A_{11}|\alpha_1^2 + 2|A_{12}|\alpha_1\alpha_2 + \ldots + |A_{kk}|\alpha_k^2)\right\}, \tag{4}$$

after fixing the constant of integration. Similarly, by starting from (4) we may proceed to the probability function (1) by solving correspondingly k simultaneous equations.

The same ideas may now be used to derive the Wishart distribution. Let $x_{1l}, x_{2l}, \ldots, x_{kl}\,(l = 1, 2, \ldots, N)$ constitute a sample from a multivariate normal population with probability function (1). Let the joint m.g.f. of the $\frac{1}{2}k(k+1)$ second-order statistics

$$v_{ii} \text{ or } \sigma_i^2 = \frac{1}{N-1}\left\{\sum_{l=1}^{N} x_{il}^2 - \left(\sum_{l=1}^{N} x_{il}\right)^2 \Big/ N\right\}$$

and

$$v_{ij} \text{ or } \rho_{ij}\sigma_i\sigma_j = \frac{1}{N-1}\left\{\sum_{l=1}^{N} x_{il}x_{jl} - \left(\sum_{l=1}^{N} x_{il}\sum_{l=1}^{N} x_{jl}\right)\Big/ N\right\},$$

be given by

$$M(\alpha_{11},\alpha_{12},\ldots,\alpha_{kk}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\phi_1\phi_2\ldots\phi_N \exp\{\alpha_{11}v_{11}+\alpha_{12}v_{12}+\ldots+\alpha_{kk}v_{kk}\}\,dx_{11}dx_{12}\ldots dx_{kN},$$

where $\phi_i \equiv \phi(x_{i1},x_{i2},\ldots,x_{ik})$, and in α_{ij} $i \leqslant j$.

Now

$$a_{i1}\frac{\partial}{\partial\alpha_{j1}}M + a_{i2}\frac{\partial}{\partial\alpha_{j2}}M + \ldots + a_{ii}\frac{\partial}{\partial\alpha_{ji}}M + \ldots + a_{ik}\frac{\partial}{\partial\alpha_{jk}}M$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\phi_1\phi_2\ldots\phi_N(a_{i1}v_{j1}+a_{i2}v_{j2}+\ldots+a_{ii}v_{ji}+\ldots+a_{ik}v_{jk})$$
$$\times\exp\{\alpha_{11}v_{11}+\ldots\}\,dx_{11}dx_{12}\ldots dx_{kN} \text{ (where } a_{pq}\equiv a_{qp} \text{ and } v_{pq}\equiv v_{qp} \text{ and } \alpha_{pq}\equiv\alpha_{qp})$$

$$= \frac{1}{N-1}\sum_{l=1}^{N}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\phi_1\ldots(\phi_l)\ldots\phi_N\left[\int_{-\infty}^{\infty}\phi_l(a_{i1}x_{l1}+a_{i2}x_{l2}+\ldots+a_{ii}x_{li}+\ldots+a_{ik}x_{lk})\right.$$
$$\left.\times\left(x_{jl}-\sum_{l=1}^{N}x_{jl}/N\right)\exp(\alpha_{11}v_{11}+\ldots)\,dx_{li}\right]dx_{11}\ldots(dx_{li})\ldots dx_{kN}.\dagger$$

Integrating the function inside the square bracket by parts with respect to x_{li}, and remembering that by equation (2)

$$\phi_l(a_{i1}x_{l1}+a_{i2}x_{l2}+\ldots+a_{ik}x_{lk}) = -\frac{\partial}{\partial x_{li}}\phi,$$

this becomes

$$\int_{-\infty}^{\infty}\phi_l\exp(\alpha_{11}v_{11}+\ldots)\left[\frac{1}{N-1}\left(x_{jl}-\sum_{l=1}^{N}x_{jl}/N\right)\left\{\alpha_{i1}\left(x_{l1}-\sum_{1}^{N}x_{l1}/N\right)+\alpha_{i2}\left(x_{l2}-\sum_{1}^{N}x_{l2}/N\right)+\ldots\right.\right.$$
$$\left.\left.+2\alpha_{ii}\left(x_{li}-\sum_{l=1}^{N}x_{li}/N\right)+\ldots+\alpha_{ik}\left(x_{lk}-\sum_{1}^{N}x_{lk}/N\right)\right\}+\frac{N-1}{N}\delta_{ij}\right]dx_{li},$$

where $\delta_{ij}\begin{cases}=1, & i=j\\ =0, & i\neq j\end{cases}$.

Therefore, after summing for all values of l, we have

$$a_{i1}\frac{\partial}{\partial\alpha_{j1}}M + a_{i2}\frac{\partial}{\partial\alpha_{j2}}M + \ldots + a_{ii}\frac{\partial}{\partial\alpha_{ji}}M + \ldots + a_{ik}\frac{\partial}{\partial\alpha_{jk}}M$$
$$= \frac{\alpha_{i1}}{N-1}\frac{\partial}{\partial\alpha_{j1}}M + \frac{\alpha_{i2}}{N-1}\frac{\partial}{\partial\alpha_{j2}}M + \ldots + 2\frac{\alpha_{ii}}{N-1}\frac{\partial}{\partial\alpha_{ji}}M + \ldots + \frac{\alpha_{ik}}{N-1}\frac{\partial}{\partial\alpha_{jk}}M + M\delta_{ij}.$$

It is thus clear that we may write down the following system of simultaneous equations:

$$\left.\begin{aligned}\left(a_{i1}-\frac{\alpha_{i1}}{N-1}\right)\frac{\partial}{\partial\alpha_{j1}}M + \left(a_{i2}-\frac{\alpha_{i2}}{N-1}\right)\frac{\partial}{\partial\alpha_{j1}}M + \ldots + \left(a_{ii}-2\frac{\alpha_{ii}}{N-1}\right)\frac{\partial}{\partial\alpha_{ji}}M + \ldots + \left(a_{jk}-\frac{\alpha_{ik}}{N-1}\right)\frac{\partial}{\partial\alpha_{jk}}M &= M \quad (i=j)\\ &= 0 \quad (i\neq j),\end{aligned}\right\} \tag{5}$$

where $i,j = 1,2,\ldots,k;\quad a_{ij}\equiv a_{ji};\quad \alpha_{ij}\equiv\alpha_{ji}$.

Solving the k equations in $\frac{1}{M}\frac{\partial}{\partial\alpha_{jr}}M, r = 1,2,\ldots,k$, obtained by giving i the k different values from 1 to k, it clearly follows that

$$\frac{1}{M}\frac{\partial}{\partial\alpha_{jr}}M = \frac{|A^*_{jr}|}{|A^*|},$$

where
$$|A^*| = \begin{vmatrix} a_{11}-\frac{2}{N-1}\alpha_{11} & a_{12}-\frac{1}{N-1}\alpha_{12} & \cdots & a_{1k}-\frac{1}{N-1}\alpha_{1k}\\ a_{12}-\frac{1}{N-1}\alpha_{12} & a_{22}-\frac{2}{N-1}\alpha_{22} & \cdots & a_{2k}-\frac{1}{N-1}\alpha_{2k}\\ \cdots & \cdots & \cdots & \cdots\\ a_{1k}-\frac{1}{N-1}\alpha_{1k} & a_{2k}-\frac{1}{N-1}\alpha_{2k} & \cdots & a_{kk}-\frac{2}{N-1}\alpha_{kk}\end{vmatrix},$$

and where $|A^*_{jr}| = |A^*_{rj}|$ is the co-factor of the jth row and rth column in $|A^*|$.

† $\phi_1\ldots(\phi_l)\ldots\phi_N$ means that ϕ_l is left out in forming the product of the other $N-1$ factors ϕ_i.

Further, since $$\frac{\partial}{\partial \alpha_{jr}} |A^*| = -\frac{2}{N-1} |A^*_{jr}|; \ (j, r = 1, 2, \ldots, k),$$

it follows that $$\log M = -\frac{N-1}{2} \log |A^*| + \log C,$$

and $M = C |A^*|^{-\frac{1}{2}(N-1)}$, where C follows immediately from $M(0, 0, \ldots, 0) = 1$. Next, to obtain the probability function $\psi \equiv \psi(v_{11}, v_{12}, \ldots, v_{kk})$, let us write down the differential equations in ψ corresponding to the system (5) above, using the relations similar to (3) above and also that $\alpha_{ir} \frac{\partial}{\partial \alpha_{jr}} M$ is the m.g.f. of

$$-\frac{\partial}{\partial v_{ir}}(v_{jr}\psi) = -v_{jr}\frac{\partial}{\partial v_{ir}}\psi \quad (i \neq j)$$

$$= -v_{jr}\frac{\partial}{\partial v_{ir}}\psi - \psi \quad (i = j).$$

Thus $$\left(v_{j1}\frac{\partial}{\partial v_{i1}}\psi + v_{j2}\frac{\partial}{\partial v_{i2}}\psi + \ldots + 2v_{ji}\frac{\partial}{\partial v_{ii}}\psi + \ldots + v_{jk}\frac{\partial}{\partial v_{ik}}\psi\right) \Big/ (N-1)$$

$$+ (a_{i1}v_{j1} + a_{i2}v_{j2} + \ldots + a_{ik}v_{jk})\psi = \begin{cases} \dfrac{N-k-2}{N-1}\psi & (i=j) \\ 0 & (i \neq j). \end{cases}$$

Clearly, writing $$|V| = \begin{vmatrix} v_{11} & v_{12} & \ldots & v_{1k} \\ v_{12} & v_{22} & \ldots & v_{2k} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ v_{1k} & v_{2k} & \ldots & v_{kk} \end{vmatrix},$$

and $|V_{ij}|$ as the co-factor of v_{ij}, we have, on solving the k equations in $\frac{1}{\psi}\frac{\partial}{\partial v_{ir}}\psi$ $(r = 1, 2, \ldots, k)$, obtained by giving j the different values from 1 to k:

$$\frac{1}{\psi}\frac{\partial}{\partial v_{ii}}\psi = \frac{(N-k-2)|V_{ii}| - (N-1)a_{ii}|V|}{2|V|},$$

where $$\frac{\partial}{\partial v_{ii}}|V| = |V_{ii}|,$$

and for $i \neq r$, $$\frac{1}{\psi}\frac{\partial}{\partial v_{ir}}\psi = \frac{(N-k-2)|V_{ir}| - (N-1)a_{ir}|V|}{|V|},$$

where $$\frac{\partial}{\partial v_{ir}}|V| = 2|V_{ir}|.$$

On integration

$$\log \psi = \tfrac{1}{2}(N-k-2)\log|V| - \tfrac{1}{2}(N-1)(a_{11}v_{11} + 2a_{12}v_{12} + \ldots + a_{kk}v_{kk}) + \log C',$$

or $$\psi = C'|V|^{\frac{1}{2}(N-k-2)} \exp\{-\tfrac{1}{2}(N-1)(a_{11}v_{11} + 2a_{12}v_{12} + \ldots + a_{kk}v_{kk})\},$$

where C' is a constant.

Quality control systems based on inaccurately measured variables

By W. M. LONG

Central Research Establishment, National Coal Board

Cases sometimes arise where the characteristic which determines the quality of a manufactured product cannot conveniently be measured directly, but instead must be estimated from the value of a related subsidiary characteristic as, for example, when the direct test of quality involves destruction of the article. The relation between the two characteristics will often be of a more or less loose nature, so that the estimation of the main characteristic from the subsidiary will be accompanied by errors, in which case we may sum up the situation by saying that the measurements of quality are inaccurate.

In the following note are discussed the measures necessary to meet this situation.

1. Let x be the estimate given by the approximate method of measuring quality of y, the true value. Assume that

(i) the regression of x on y is linear;

(ii) the successive observations x from the production system, with their associated y's, generate the bivariate normal distribution

$$\frac{1}{2\pi\sigma}\exp-\frac{1}{2\sigma^2}\{x^2+y^2-2\sqrt{(1-\sigma^2)}xy\}\,dx\,dy,$$

the variates x and y being measured in terms of their standard deviations from their mean values for the system; σ^2 is the residual variance of the regression of x on y and is equal to $(1-\rho^2)$, where ρ is the coefficient of correlation between x and y.

The conditional distribution of x, given y, is then

$$N\{\sqrt{(1-\sigma^2)}\,y, \sigma^2\}. \tag{1}$$

Clearly in the bivariate system any change in the marginal distribution of the y's must be accompanied by a corresponding change in that of the x's, *provided that the distribution* (1) *remains unaffected*. We may therefore use the marginal distribution of the x's as a check on the stability of the production system; i.e. we may base the quality control system on the x's, precisely as if they were direct measurements of quality.

2. It is not of course to be expected that a control system based on the x's will be as efficient as one based on the y's. To see how the efficiency is reduced we may consider the respective powers of the x-tests and y-tests with regard to a set of simple hypotheses. Instead of considering the number of defectives in the test sample, as is usual, we shall for convenience deal with the displacement of the sample mean from the hypothetical mean for the system; the principle of the argument remains unaffected.

Let the acceptance region (or one of them) in the test be the interval $\left(-\frac{x_0}{\sqrt{n}}, \frac{x_0}{\sqrt{n}}\right)$, where

$$\frac{1}{\sqrt{(2\pi)}}\int_{-x_0}^{x_0} e^{-\frac{1}{2}x^2}dx = (1-\alpha), \text{ say},$$

the test being based on the mean of n observations, so that if a sample mean $\bar{x}$ falls outside the above interval a change is assumed to have occurred in the production system and action is taken.

Suppose a change has occurred, in that the average level of quality has shifted from 0 to ξ, the variances of x and y, and the conditional distribution (1) remaining unchanged. This implies a shift in the x-mean from 0 to $\sqrt{(1-\sigma^2)}\,\xi = \rho\xi$, the correlation coefficient remaining unchanged.

The power of the x-test in detecting the change is

$$1-\sqrt{\frac{n}{2\pi}}\int_{-x_0/\sqrt{n}}^{x_0/\sqrt{n}} e^{-\frac{1}{2}n(\bar{x}-\rho\xi)^2}d\bar{x} = 1-\frac{1}{\sqrt{(2\pi)}}\int_{-x_0-\sqrt{n}\rho\xi}^{x_0-\sqrt{n}\rho\xi} e^{-\frac{1}{2}x^2}dx.$$

The effectiveness of the x-test, compared with that of a test based on exact measurements, may be described by the ratio

$$\left\{1-\frac{1}{\sqrt{(2\pi)}}\int_{-x_0-\sqrt{n}\rho\xi}^{x_0-\sqrt{n}\rho\xi} e^{-\frac{1}{2}x^2}dx\right\}\left\{1-\frac{1}{\sqrt{(2\pi)}}\int_{-x_0-\sqrt{n}\,\xi}^{x_0-\sqrt{n}\,\xi} e^{-\frac{1}{2}x^2}dx\right\}^{-1},$$

the denominator being the power of the test when $\rho = 1$, i.e. when the measurements of quality are exact. We shall call this ratio the 'power factor' of the x-test; the analogy with the definition of the efficiency of an estimator is obvious. The power factor is in this case a function of ρ, n, α and ξ.

In Figs. 1 and 2 the full lines show how the power factor varies with the quantity $\sqrt{(n\,\xi)}$, for $\alpha = 0{\cdot}01$ and $0{\cdot}001$ and $\rho = 0{\cdot}6$, $0{\cdot}7$, $0{\cdot}8$ and $0{\cdot}9$. It is unity at $\xi = 0$, $+\infty$ and $-\infty$ (this follows from the definition), and for values of ρ in the range considered has minima near $\sqrt{n}\,\xi = 2$, i.e. $\xi = 2/\sqrt{n}$.

The practical significance of this may be gauged from the dotted lines on the diagrams, which show how the power of the exact test (case $\rho = 1$) varies with $\sqrt{n}\,\xi$. Clearly the value of the power factor, for $\sqrt{n}\,\xi$ less than 2, say, is not of great interest, since the power of the exact test itself is then low.

The power lost through the inaccuracy of the measurements may be restored, at least partially, by increasing the size of the test sample. For the power of a test based on n_1 exact measurements is the same as that of one based on n_2 inaccurate measurements if $\sqrt{n_2}\,\rho = \sqrt{n_1}$, i.e. if $n_2 = n_1/\rho^2$. It will not of course always be possible to recover all the lost power in this way, since n_2 is limited in size, the limit being reached when all the output is being examined.

3. The joint distribution of x and y depends on five parameters, the means and variances of x and y respectively and the correlation coefficient ρ. So far in our control system we have considered one only, the x-mean; a second, the variance of x, may be covered by the ordinary quality control methods which include a check on the variance of the control variable based on the sample range. This leaves three parameters, and it will be seen that these are covered by our assumption that the distribution (1) remains unchanged.

This assumption will often be a reasonable one to make, since the relation between x and y will usually be a consequence of the fundamental natural properties of the material, properties which may be quite stable and insensitive to changes in any particular system of production.

In some cases, however, it will be advisable to carry out periodical checks on the stability of the distribution (1). This will require special samples in which both x and y are measured. The following is an appropriate significance test.

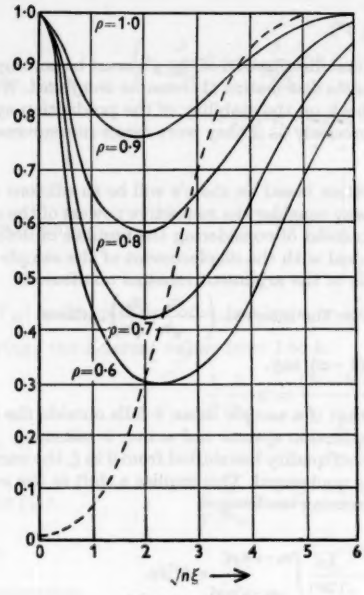


Fig. 1. $\alpha = 0{\cdot}01$.

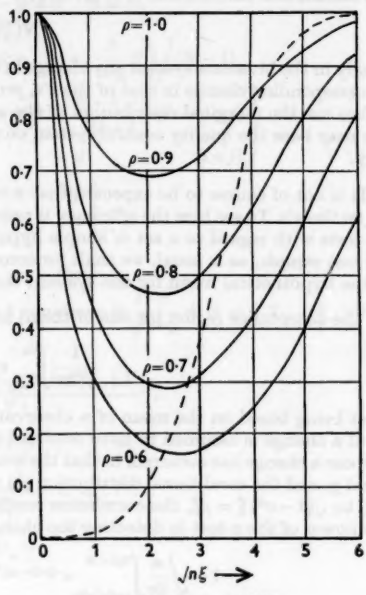


Fig. 2. $\alpha = 0{\cdot}001$.

Let (x_i, y_i) be a typical pair of observations from the test sample, there being n pairs in all, say. If the regression system is stable then $S = \sum_{i=1}^{n} (x_i - \rho y_i)^2/\sigma^2$ is distributed as χ^2 with n degrees of freedom. Let $(x - \bar{x}) = b(y - \bar{y})$ be the line of best fit for the sample. Then

$$\Sigma(x_i - \rho y_i)^2/\sigma^2 = \Sigma[x_i - \bar{x} - b(y_i - \bar{y})]^2/\sigma^2 + n(\bar{x} - \rho\bar{y})^2/\sigma^2 + (b - \rho)^2 \Sigma(y_i - \bar{y})^2/\sigma^2$$

$$= \quad S_1 \quad + \quad S_2 \quad + \quad S_3, \text{ say.}$$

The quantities S_1, S_2 and S_3 are, under the assumption of regression stability, independently distributed as χ^2 with degrees of freedom $(n-2)$, 1 and 1 respectively. By entering the values obtained in the χ^2 table with the appropriate degrees of freedom they will therefore provide tests for changes in the residual variance and the two parameters of the regression line, respectively.

We should also test S itself, for S may attain a significant value even though each of its components fails to do so. In this we should still conclude that something was amiss, without being able to specify exactly where among the various parameters a change had taken place.

A little care is necessary in judging the results; S_1 should be tested first, for if a change in the residual variance has occurred the tests for the other two parameters no longer apply. In this case the appropriate

tests for the line parameters are the F-ratios $F_{1,n-2} = (n-2)\,S_2/S_1$ and $F_{1,n-2} = (n-2)\,S_3/S_1$, where the affixes have their usual meanings in terms of degrees of freedom. A collective test

$$F_{2,n-2} = (n-2)\,(S_2+S_3)/(2S_1)$$

may also be applied, for reasons similar to those which lead us to test S in the above.

4. From the above discussion of the power factor it will be apparent that there is nothing sacrosanct, for quality control purposes, in direct measurements of quality, even though to obtain them may be perfectly practicable. The same power in testing may often be obtained through the use of larger numbers of indirect measurements, and the latter procedure will sometimes be preferable on the grounds of convenience and cheapness.

The bivariate normal integral

By J. H. CADWELL

1. In order to evaluate the bivariate integral over a rectangle with sides parallel to the axes, we can use the function

$$M(h,k,\rho) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\int_h^\infty\int_k^\infty \exp -\frac{(x^2-2\rho xy+y^2)}{2(1-\rho^2)}\,dx\,dy.$$

Tables will be found in *Tables for Statisticians and Biometricians*, volume 2 (Pearson, K. (1931)). Linear interpolation is adequate for small values of ρ. However, for $|\rho| > 0{\cdot}5$, second differences in ρ are needed for four-figure accuracy.

If the rectangle R is bounded by the lines

$$x = h_1, \quad x = h_2, \quad y = k_1, \quad y = k_2,$$

the integral over R will be equal to

$$M(h_1,k_1,\rho)+M(h_2,k_2,\rho)-M(h_1,k_2,\rho)-M(h_2,k_1,\rho).$$

The following results are needed for negative h and k:

$$\begin{aligned} M(h,k,\rho) &= \tfrac{1}{4} - \Phi(h) - M(h,-k,-\rho) \\ &= M(-h,-k,\rho) + \Phi(-h) + \Phi(-k), \end{aligned}$$

where $\Phi(x)$ is the normal integral from 0 to x.

2. We show below that the bivariate integral over R, or, indeed, over any polygon, can be evaluated by using the function

$$V(h,q) = \frac{1}{2\pi}\int_0^h\int_0^{qx/h} \exp -\tfrac{1}{2}(x^2+y^2)\,dx\,dy.$$

As this depends on only two variables, a much smaller table will suffice. Such a table, giving 6 decimal places for $V(h,q)$, has been prepared by Nicholson (1943) for arguments $h = 0{\cdot}1(0{\cdot}1)\,3{\cdot}0$ and $q = 0{\cdot}1(0{\cdot}1)\,3{\cdot}0$. Should this table not be available, the approximation derived below will give three-place accuracy for all values of the variables.

Nicholson also found an asymptotic formula; this will give results of high accuracy for large h and q. An alternative derivation is given below.

We see that $V(h,q)$ is the bivariate integral with $\rho = 0$, taken over the triangular region defined by the points $(0,0)$, $(h,0)$ and (h,q). By virtue of circular symmetry, the integral will be unaffected by rotation of this triangle about its vertex at $(0,0)$. The result

$$\Phi(h)\,\Phi(k) = V(h,k) + V(k,h)$$

follows readily. On putting $h = k$, and using the approximation for V, an improved version of the result

$$\Phi(x) \doteq \frac{1}{2}\left\{1-\exp\left(-\frac{2x^2}{\pi}\right)\right\}^{\frac{1}{2}}$$

is obtained.

3. In order to evaluate the integral ($\rho \neq 0$) over R we proceed as follows. Using the transformation

$$w = x, \quad z = (y - \rho x)/(1-\rho^2)^{\frac{1}{2}},$$

the integral over the shaded triangle is found to be

$$| V(h_1, q_{11}) - V(h_1, q_{21}) |, \quad q_{i1} = \frac{k_i - \rho h_1}{(1-\rho^2)^{\frac{1}{2}}} \quad (i = 1, 2).$$

A similar transformation gives for the triangle OAB

$$| V(k_1, p_{11}) - V(k_1, p_{21}) |, \quad p_{i1} = \frac{h_i - \rho k_1}{(1-\rho^2)^{\frac{1}{2}}} \quad (i = 1, 2).$$

For negative values we need the results

$$V(h, k) = -V(h, -k) = -V(-h, k) = V(-h, -k).$$

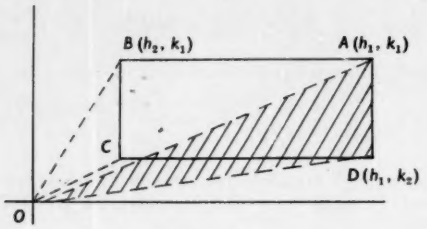


Fig. 1

In the above figure the two integrals are to be added, and two similar integrals taken from their sum. When the origin lies inside R, all four integrals will be added.

With the obvious definitions of p_{ij}, etc., the integral over R is

$$| V(h_1, q_{11}) + V(h_2, q_{22}) + V(k_1, p_{11}) + V(k_2, p_{22}) - V(h_1, q_{21}) - V(h_2, q_{12}) - V(k_1, p_{21}) - V(k_2, p_{12}) |.$$

However, it is more convenient to evaluate the four portions separately, and then combine them suitably.

4. Using the tables for $V(h, q)$ we can evaluate the general integral over any polygon. A rotation of axes followed by a change of scale will reduce the problem to that of evaluating

$$\frac{1}{2\pi} \iint_P \exp -\tfrac{1}{2}(x^2 + y^2)\, dx\, dy.$$

This integral is the sum of a number of integrals over areas of the type shown shaded in Fig. 2.

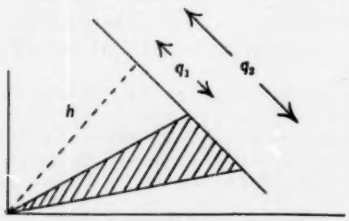


Fig. 2

Because of the circular symmetry, this area will contribute

$$V(h, q_2) - V(h, q_1).$$

Thus the problem is solved.

5. In order to obtain an approximate formula for V, we replace the triangular area by the sector of a circle shown in Fig. 3. The integral over the sector is

$$\frac{\alpha}{2\pi}\{1-\exp(-\tfrac{1}{2}R^2)\}.$$

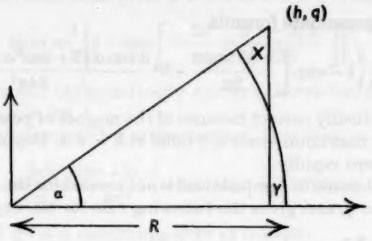


Fig. 3

Throughout X the integrand is smaller than at any point of Y. Thus the integral over the triangle will be smaller than that over the circular sector. However, if α is not too large, we assume that the integrand has the same constant value over X and Y. Then we choose R to make X and Y have equal areas. This amounts to making the triangle and sector have the same area and gives the approximation

$$V(h,q) \doteq \frac{\tan^{-1} q/h}{2\pi}\left\{1-\exp\frac{-hq}{2\tan^{-1} q/h}\right\}.$$

For values of $q>h$ we evaluate $V(q,h)$ and use the result

$$V(h,q)+V(q,h) = \Phi(h)\,\Phi(q).$$

The error of this approximation, after a change of sign, is

$$\frac{1}{2\pi}\int_0^{\alpha}\exp\frac{-h^2}{2\cos^2\theta}\,d\theta-\frac{\alpha}{2\pi}\left(1-\exp\frac{-h^2\tan\alpha}{2\alpha}\right).$$

Differentiating with regard to α and h we find that it is a maximum (in the region $q \leqslant h$) for

$$\alpha=\frac{\pi}{4},\quad \frac{h}{\sqrt{(2\pi)}}\exp-h^2\left(\frac{2}{\pi}-\frac{1}{2}\right) = \Phi(h).$$

An approximate solution gives $h = 1{\cdot}76$, with a maximum error of 0·0015. Fig. 4 shows the error graphs for $q = h$ and $q = 0{\cdot}75h$. It is evident that, in general, errors will be much less than the maximum value. The value given by the approximation is always too large.

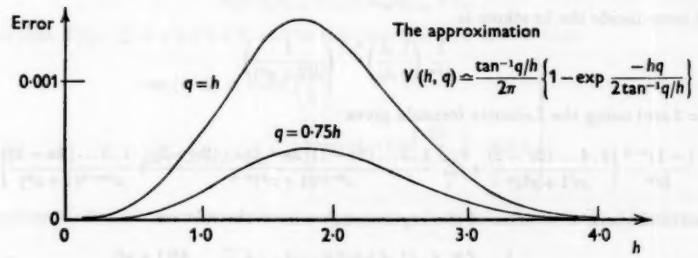


Fig. 4

6. To improve on this approximation we expand the integrand about R as

$$\left(1+\frac{R^2-r^2}{2}\right)\exp\frac{-R^2}{2}.$$

Now we choose R so that this quantity gives equal and opposite values when integrated over X and Y. This gives the following result:

$$4hq(1+\tfrac{1}{2}R^2)-hq(h^2+\tfrac{1}{3}q^2)=4R^2\alpha(1+\tfrac{1}{2}R^2)-\alpha R^4.$$

Assuming that $$R^2=hq/\alpha+\epsilon$$

with ϵ small, we derive the approximate formula

$$V(h,q)\simeq\frac{\alpha}{2\pi}\left\{1-\exp\left[\frac{-h^2\tan\alpha}{2\alpha}+h^4\frac{\alpha\tan\alpha(3+\tan^2\alpha)-3\tan^2\alpha}{24\alpha^2}\right]\right\}.$$

This formula is not asymptotically correct because of the neglect of powers of ϵ in deriving it. However, over the range, $0<h<3{\cdot}5$, the maximum error is $0{\cdot}0003$ at $h=2{\cdot}4$. Beyond $3{\cdot}5$ the simpler formula gives a smaller error that tends to zero rapidly.

The correction term is troublesome to compute and is not needed for the small q. This suggests expanding the correcting factor about $h=q$, and gives the following rule for the exponential index:

If $\left.\begin{matrix}h<3{\cdot}5\\0{\cdot}75h<q<h\end{matrix}\right\}$ use $\dfrac{-hq}{2\tan^{-1}q/h}+0{\cdot}04h^3q-0{\cdot}03h^4,$

otherwise the simpler form is to be used. This procedure will never give an error greater than $0{\cdot}0005$, and, in general, will be much nearer the true value.

7. We now derive an asymptotic expansion due to Nicholson (1943) for the integral

$$I(h,q)=\frac{1}{2\pi}\int_h^\infty\int_{qx/h}^\infty\exp-\frac{(x^2+y^2)}{2}\,dx\,dy.$$

By considering integrals over appropriate regions we see that

$$I(h,q)+\tfrac{1}{2}\Phi(h)=V(h,q)+\frac{1}{2\pi}\cot^{-1}\frac{q}{h}.$$

Transforming to polars and using an appropriate change of variables,

$$I(h,q)=\frac{1}{2\pi}\int_\alpha^{\frac{1}{2}\pi}\exp-\frac{h^2}{2\cos^2\theta}\,d\theta=\frac{1}{2\pi}\int_x^\infty\frac{1}{(1+t^2)}\exp\frac{-h^2(1+t^2)}{2}\,dt.$$

Here α and x are defined by $$x=\tan\alpha=\frac{q}{h}.$$

Repeated integration by parts then gives

$$I\sim\frac{1}{2\pi}\exp\frac{-h^2(1+x^2)}{2}\left\{\frac{1}{h^2x(1+x^2)}-\frac{1}{h^4}\left(\frac{1}{x^3(1+x^2)}+\frac{2}{x(1+x^2)^2}\right)+\ldots\right\}.$$

The general term inside the brackets is

$$\frac{1}{h^{2n}}\left(\frac{1}{x}\frac{d}{dx}\right)^{n-1}\left(\frac{1}{x(1+x^2)}\right).$$

Putting $x^2=t$ and using the Leibnitz formula gives

$$\frac{(-1)^{n-1}}{h^{2n}}\left\{\frac{2.4\ldots(2n-2)}{x(1+x^2)^n}+\sum_1^{n-2}\frac{1.3\ldots(2r-1)(2r+2)\ldots(2n-2)}{x^{2r+1}(1+x^2)^{n-r}}+\frac{1.3\ldots(2n-3)}{x^{2n-1}(1+x^2)}\right\}.$$

If S_n denotes the sum of n terms of the expansion, we have the relation

$$I=S_n+\frac{1}{2\pi h^{2n}}\int_x^\infty\frac{d}{dx}\left(\frac{1}{x}\frac{d}{dx}\right)^{n-1}\left\{\frac{1}{x(1+x^2)}\right\}\exp-\frac{h^2(1+x^2)}{2}\,dx.$$

From the formula for the nth term we see that this integral is of alternating sign. Thus, S_n will be alternately too big and too small, since the integral will take alternate positive and negative values. Beyond a point depending on h and x, the terms begin to increase. We thus have the rule, that if the rth term be smallest, I will lie between S_{r-1} and S_r. Further terms cannot improve on this information.

8. The well-known approximation discussed by Pólya (1949),

$$\Phi(x) \simeq \frac{1}{2}\left(1-\exp\frac{-2x^2}{\pi}\right)^{\frac{1}{2}},$$

arises from the formula of §5. Its maximum error is 0·003 at $x = 1{\cdot}6$. Using the improved formula of §6

$$\Phi(x) \simeq \frac{1}{2}\left\{1-\exp\left[-\frac{2x^2}{\pi}+\frac{2(\pi-3)x^4}{3\pi^2}\right]\right\}^{\frac{1}{2}}.$$

We note that this formula is not asymptotically correct. However, over the range $0<x<3{\cdot}5$, the maximum error 0·0007 arises at $x = 2{\cdot}5$.

While the method used could be carried a stage further, it is easy to make empirical adjustments. The formula

$$-\frac{2x^2}{\pi}+\frac{2(\pi-3)x^4}{3\pi^2}-0{\cdot}0005x^6+0{\cdot}00002x^8$$

for the index of the exponential gives a maximum error of 0·00005.

Acknowledgement is made to the Chief Scientist, Ministry of Supply, for permission to publish this paper.

REFERENCES

Nicholson, C. (1943). *Biometrika*, **33**, 59.

Pearson, K. (Ed.) (1931). *Tables for Statisticians and Biometricians*, **2**. Biometrika Office, University College, London.

Pólya, G. (1949). *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, p. 63. University of California Press.

Note on the variate differences of autoregressive series

By BENTO MURTEIRA

1. The object of the present note is to give a simple expression for

$$\lim_{k\to\infty}\frac{\operatorname{var}(\Delta^k x_t)}{\binom{2k}{k}} = \lim_{k\to\infty} V_k, \tag{1}$$

when x_t is a series generated by an autoregressive scheme of order m,

$$x_t+\alpha_1 x_{t-1}+\ldots+\alpha_m x_{t-m} = \epsilon_t. \tag{2}$$

We assume that $E\{\epsilon_t\epsilon_{t+v}\} = 0$ for $v \neq 0$, and in this case it is well known that

$$\left.\begin{aligned}\operatorname{var}(\Delta^k\epsilon_t) &= \operatorname{var}\epsilon\binom{2k}{k}\\ \operatorname{cov}(\Delta^k\epsilon_t\Delta^k\epsilon_{t+s}) &= \begin{cases}\operatorname{var}\epsilon(-1)^s\binom{2k}{k-s} & k\geqslant s,\\ 0 & k<s.\end{cases}\end{aligned}\right\} \tag{3}$$

Let us consider for fixed k the autocovariance generating function of the series $\Delta^k x_t$ given by

$$\begin{aligned}\overline{G_k}(z) &= G(z)\,G_k(z)\\ &= \sum_{u=-\infty}^{\infty} C_u z^u \sum_{v=-\infty}^{\infty} C_v^{(k)} z^v\\ &= \sum_{s=-\infty}^{\infty} \overline{C}_s^{(k)} z^s,\end{aligned} \tag{4}$$

where $$G(z) = \frac{1}{(1+\alpha_1 z+\ldots+\alpha_m z^m)}\,\frac{1}{(1+\alpha_1 z^{-1}+\ldots+\alpha_m z^{-m})},$$

$$C_s^{(k)} = E\{\Delta^k \epsilon_t \Delta^k \epsilon_{t+s}\},$$

$$\overline{C}_s^{(k)} = E\{\Delta^k x_t \Delta^k x_{t+s}\}$$

and $$\operatorname{var}(\Delta^k x_t) = \overline{C}_0^{(k)}.$$

The expansion of $G(z)$ as a power series in z is to be regarded as a purely formal one determining the C's in terms of α's. It will, however, converge for $z = \pm 1$ if the series is stationary, for then all roots of the equation $t^m+\alpha_1 t^{m-1}+\ldots+\alpha_m = 0$ are less than unity in modulus.

We can see easily that

$$\overline{C}_0^{(k)} = \sum_{s=-\infty}^{\infty} C_s C_{-s}^{(k)}$$

$$= \sum_{s=-\infty}^{\infty} C_s C_s^{(k)},$$

and

$$\lim_{k\to\infty} \frac{\overline{C}_0^{(k)}}{\binom{2k}{k}} = \operatorname{var}\varepsilon \sum_{s=-\infty}^{\infty} C_s(-1)^s \lim_{k\to\infty} \frac{\binom{2k}{k-s}}{\binom{2k}{k}}$$

$$= \operatorname{var}\varepsilon \sum_{s=-\infty}^{\infty} C_s(-1)^s$$

$$= \operatorname{var}\varepsilon\, G(-1)$$

$$= \operatorname{var}\varepsilon\,[1-\alpha_1+\alpha_2-\ldots+(-1)^m\alpha_m]^{-2}. \qquad (5)$$

2. This result was applied to Kendall's (1946) experimental series:

	Series 1	Series 2	Series 3	Series 4
V_0	2535·11	3414·42	3900·94	2001·20
V_1	604·10	523·87	1472·92	3377·47
V_2	278·79	239·46	523·63	3958·74
...	...	...	...	...
V_{20}	102·14	133·23	101·51	3415·88
Limit (equation (5))	130·01	112·56	104·62	3335·33

The agreement seems fair if we note that our result is for an infinite series. It is interesting to see that for series 1 the expected value of V_{20} is 137·66.

It is also of some interest to note that for a 'wandering' series, when the α's, except α_1, are zero and $\alpha_1 = -1$, the limit (5) is infinite; but that for $\alpha_1 < -1$ the limit is finite. In the latter case, however, the series is not stationary and the result fails to hold.

3. The generalization to autoregressive schemes with moving average errors such as that below,

$$x_t+\alpha_1 x_{t-1}+\ldots+\alpha_m x_{t-m} = \epsilon_t+\beta_1\epsilon_{t-1}+\ldots+\beta_n\epsilon_{t-n}, \qquad (6)$$

is straightforward. The formula (4) still holds,

$$\overline{G}_k(z) = G(z)\,G_k(z),$$

but now $$G(z) = \frac{1+\beta_1 z+\ldots+\beta_n z^n}{1+\alpha_1 z+\ldots+\alpha_m z^m}\,\frac{1+\beta_1 z^{-1}+\ldots+\beta_n z^{-n}}{1+\alpha_1 z^{-1}+\ldots+\alpha_m z^{-m}}.$$

The same reasoning as above leads to the conclusion that

$$\lim_{k\to\infty} V_k = \operatorname{var}\varepsilon\, G(-1)$$

$$= \operatorname{var}\varepsilon\left[\frac{1-\beta_1+\beta_2-\ldots+(-1)^n\beta_n}{1-\alpha_1+\alpha_2-\ldots+(-1)^m\alpha_m}\right]^2. \qquad (7)$$

REFERENCE

KENDALL, M. G. (1946). *Contributions to the Study of Oscillatory Time-Series.* Cambridge University Press.

Note on the inversion theorem

By J. GIL-PELAEZ

If $F(x)$ is a one-dimensional distribution function, its corresponding characteristic function is the complex function of the real variable t:

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x).$$

The theorem which proves that the distribution is uniquely determined by the characteristic function is called the inversion theorem. This was first demonstrated by Lévy, who found the expression

$$F(x) - F(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1-e^{-itx}}{it}\phi(t)\, dt.$$

The constant $F(0)$ is no obstacle to an ascertainment of the distribution, because by subtraction we have for every interval (a, b)

$$\int_a^b dF(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ita}-e^{-itb}}{it}\phi(t)\, dt.$$

Furthermore, when the distribution belongs to the continuous type the constant is also eliminated by differentiation,

$$f(x) = F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\, dt.$$

This is sufficient in the majority of cases, but when we are interested in the explicit expression of the distribution function, the constant $F(0)$ is a substantial inconvenience. I therefore present a derivation of the inversion theorem by a new method which has the advantage of evaluating the constant term. I shall, in fact, show that

$$F(x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^{\infty} \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{it}\, dt.$$

In order to prove this, let us consider the step-function

$$\operatorname{sign}(y-x) = \frac{2}{\pi}\int_0^{\infty} \frac{\sin t(y-x)}{t}\, dt = \begin{cases} -1 & \text{if } y < x, \\ 0 & \text{if } y = x, \\ 1 & \text{if } y > x, \end{cases}$$

and observe that

$$\int_{-\infty}^{\infty} \operatorname{sign}(y-x)\, dF(y) = 1 - 2F(x).$$

Now, for any positive numbers ϵ, λ, we have

$$\frac{1}{\pi}\int_{\epsilon}^{\lambda} \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{it}\, dt = \frac{1}{\pi}\int_{\epsilon}^{\lambda}\int_{-\infty}^{\infty} \frac{e^{-it(y-x)} - e^{it(y-x)}}{it}\, dF(y)\, dt$$

$$= -\frac{2}{\pi}\int_{\epsilon}^{\lambda}\int_{-\infty}^{\infty} \frac{\sin t(y-x)}{t}\, dF(y)\, dt = -\int_{-\infty}^{\infty} dF(y)\, \frac{2}{\pi}\int_{\epsilon}^{\lambda} \frac{\sin t(y-x)}{t}\, dt.$$

The inversion of the order of integration has been possible because for $\epsilon < t < \lambda$ and $-\infty < y < \infty$

$$\left| \frac{\sin t(y-x)}{t} \right| < \frac{1}{\epsilon}.$$

When ϵ tends to zero and λ tends to infinity, we have

$$\frac{1}{\pi}\int_0^{\infty} \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{it}\, dt = -\int_{-\infty}^{\infty} dF(y) \lim \frac{2}{\pi}\int_{\epsilon}^{\lambda} \frac{\sin t(y-x)}{t}\, dt$$

$$= -\int_{-\infty}^{\infty} \operatorname{sign}(y-x)\, dF(y) = 2F(x) - 1,$$

which establishes the theorem. We observe that it is possible to take the limit inside the integral sign because

$$\int_{\epsilon}^{\lambda} \frac{\sin t(y-x)}{t}\, dt$$

is a continuous function of ϵ, λ with a bounded modulus.

When the characteristic function is of the following kind,

$$\phi(t) = \sum_r p_r e^{tix_r},$$

the formula of the inversion theorem gives us the following function:

$$F(x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{\sum_r p_r e^{it(x-x_r)} - \sum_r p_r e^{-it(x-x_r)}}{it}\,dt$$

$$= \frac{1}{2} + \sum_r p_r \frac{1}{\pi}\int_0^\infty \frac{\sin t(x-x_r)}{t}\,dt = \frac{1}{2} + \frac{1}{2}\sum_r p_r \operatorname{sign}(x-x_r).$$

This is a step-function with a saltus p_r at x_r. We must observe that at the points of discontinuity the value of the function is the half-sum of the limits on the right and on the left. This is not the value if the distribution function is taken to be continuous on the left, but depends on the Dirichlet integral which has been used in the proof.

I have to thank Mr D. G. Kendall for his helpful suggestions.

A note on the statistical independence of quadratic forms in the analysis of variance

By J. A. NELDER, *National Vegetable Research Station*

A widely used notation in writing out the algebraic identities in the analysis of variance is that employing multiple suffixes and using dots to denote averaging over the suffices concerned. Thus with $p \times q$ two-way classification (to take a simple example) we have

$$\sum_{i,j} x_{ij}^2 \equiv \sum_{i,j} x_{..}^2 + \sum_{i,j}(x_{i.}-x_{..})^2 + \sum_{i,j}(x_{.j}-x_{..})^2 + \sum_{i,j}(x_{ij}-x_{i.}-x_{.j}+x_{..})^2. \tag{1}$$

Let us denote the four linear forms in the x_{ij}, namely,

$$x_{..}, \quad x_{i.}-x_{..}, \quad x_{.j}-x_{..} \quad \text{and} \quad x_{ij}-x_{i.}-x_{.j}+x_{..}$$

by

$$y_{1ij}, \quad y_{2ij}, \quad y_{3ij} \quad \text{and} \quad y_{4ij}$$

respectively. Then writing x, y_1, y_2, y_3, y_4 for the five $pq \times 1$ vectors

$$\{x_{ij}\}, \quad \{y_{1ij}\}, \quad \{y_{2ij}\}, \quad \{y_{3ij}\}, \quad \{y_{4ij}\}$$

and putting $y_\alpha = B_\alpha x$ we note that

(i) B_α is symmetrical for all α,

(ii) $B_1 + B_2 + B_3 + B_4 = I$, since

$$x_{ij} \equiv x_{..} + (x_{i.}-x_{..}) + (x_{.j}-x_{..}) + (x_{ij}-x_{i.}-x_{.j}+x_{..}),$$

(iii) $B_\alpha B_\beta$ is skew symmetric, $\alpha \neq \beta$, since each cross-product such as

$$\sum_{i,j}(x_{.j}-x_{..})(x_{ij}-x_{i.}-x_{.j}+x_{..}) = x'B_3B_4x$$

vanishes identically in the x_{ij}.

Now, although accounts of the analysis of variance note the separate vanishing of each of the cross-products in (1) and similar identities, it does not seem to have been pointed out that when the x_{ij} are independent normal variates with zero mean and unit variance, the conditions (i), (ii) and (iii) above on the B matrices are sufficient for the quadratic forms on the right-hand side of (i), i.e.

$$(B_\alpha x)' B_\alpha x = x'B_\alpha' B_\alpha x = x'B_\alpha^2 x \quad (\alpha = 1, 2, 3, 4),$$

to be independently distributed as χ^2 and that the number of degrees of freedom associated with the form $x'B_\alpha^2 x$ is trace B_α. Thus the number of degrees of freedom of $\sum_{i,j}(x_{ij}-x_{i.}-x_{.j}+x_{..})^2$ is given by

$$\text{trace } B_4 = pq\left(1 - \frac{1}{q} - \frac{1}{p} + \frac{1}{pq}\right) = pq - p - q + 1,$$

a familiar result.

Since this approach leads to easily verifiable sufficient conditions for the independence of the quadratic forms in the analysis of variance when expressed in the above notation, and to a simple method of evaluating the number of degrees of freedom associated with each such form, it may be worth giving a proof.

The result follows from the following theorem in matrix algebra, contained implicitly in a paper of A. T. Craig (1943):

If

(*a*) $A_1, \ldots, A_n$ are n real symmetric $N \times N$ matrices,

(*b*) $A_1 + A_2 + \ldots + A_n = I$,

(*c*) $A_i A_j = 0$ $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n;\ i \neq j)$,

then the A_i are idempotent and there exists an orthogonal matrix H such that non-zero parts of $H' A_i H$ are disjunct diagonal submatrices with unit diagonal elements. (It may be noted that this is Cochran's Theorem (Cochran, 1934) as given in matrix notation by Aitken (1950) except that we have substituted condition (*c*) for the condition that the sum of the ranks of the A_i equals N. The two conditions are in fact equivalent.) If x represents an $N \times 1$ matrix of independent normal variates with zero mean and unit variance, then by standard normal theory the quadratic forms $x' A_i x (= x' A_i^2 x$, since the A's are idempotent) are independently distributed as χ^2 with degrees of freedom given by trace $H' A_i H$ = trace A_i, since H is orthogonal.

The B matrices, however, in identity (1) are not immediately seen to satisfy condition (*c*) of the theorem. In fact, the skew-symmetry condition (iii), in conjunction with conditions (i) and (ii), implies this as will now be shown.

Pre- and post-multiply by A_1 the relation

$$A_1 + A_2 + \ldots + A_n = I$$

and add giving $$2A_1^2 + (A_1 A_2 + A_2 A_1) + \ldots + (A_1 A_n + A_n A_1) = 2A_1;$$

from (i) and (iii) $$A_1 A_2 + A_2 A_1 = \ldots = A_1 A_n + A_n A_1 = 0,$$

therefore $A_1^2 = A_1$ and generally $A_i^2 = A_i$ $(i = 1, 2, \ldots, n)$.

Now since $$A_i A_j + A_j A_i = 0,$$

therefore $$A_i^2 A_j + A_i A_j A_i = 0,$$

and so $$A_i^2 A_j = A_i A_j = -A_i A_j A_i,$$

which is symmetric. Hence $A_i A_j$ is both symmetric and skew symmetric and so is equal to the zero matrix. Thus in conjunction with conditions (i) and (ii), condition (iii) implies condition (*c*) and the result follows.

REFERENCES

Aitken, A. C. (1950). On the statistical independence of quadratic forms in normal variates. *Biometrika*, **37**, 93–6.

Cochran, W. G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. *Proc. Camb. Phil. Soc.* **30**, 178–91.

Craig, A. T. (1943). Note on the independence of certain quadratic forms. *Ann. Math. Statist.* **14**, 195–7.

REVIEWS

Statistics, volume II. By N. L. JOHNSON and H. TETLEY. xi + 318 pp. Cambridge University Press. 1950. Price 20*s*.

This volume represents the completion of the new text-book designed to cover the statistical requirements of Parts I and II of the syllabus of the Institute of Actuaries Examinations. The first part, which was recently reviewed in *Biometrika*, vol. 37, p. 453, dealt with the methods of descriptive statistics, questions of statistical inference and the simpler statistical tests. This second volume deals with the more exact theory of sampling, statistical tests based thereon, including the general theory of statistical tests, analysis of variance, correlation analysis, curve fitting and graduation.

In preparing this second volume the authors have had to keep in mind the mathematical equipment of the students for whom the book has been primarily prepared. As this precludes the use of complex variable analysis they have been unable to utilize the various elegant derivations and proofs which follow from the use of characteristic functions. Similarly, a significant omission from the book is any reference to generating functions and moment-generating functions in particular. Whilst the omission of these and other special techniques does not prevent a successful mathematical development, some direct reference to them and their application would have added to the presentation.

The chapter numbers follow on from the first volume and Chapter 11, which is appropriately headed the calculus of distribution functions, is based on random sampling from a normal population and develops the theory of the commoner distributions, including χ^2_f, t_f and F_{f_1,f_2}. The practical application of these distributions in deriving significance levels instead of the approximate methods adopted in Chapter 10 is developed in Chapter 12.

Chapter 13 extends the earlier work on the binomial distribution to the multinomial distribution. This leads naturally to the problem of ranked variables and to the distribution of the range and of the median. The concept of power is introduced in Chapter 14 with a discussion on statistical tests and the meaning of likelihood.

After a short discussion on the various principal sampling systems the essential features of analysis of variance are dealt with in Chapter 15 and the principles of correlation analysis follow in the next chapter.

The final chapter is devoted to the general principles of curve fitting and graduation, including the question of goodness of fit.

As in the first volume each chapter concludes with a number of examples, many of which cover important results which could not be included in the text. They thus vary in difficulty between fairly wide limits. Some mathematical points have been relegated to appendices, which also include tables of χ^2, t and F. A very valuable feature is the provision at the end of each chapter of a fairly comprehensive bibliography related to the subject-matter of the chapter. This inevitably involves a repetition of titles, but will be of considerable value to the student who requires more extended treatment of a particular topic.

With the completion of this second volume it is now possible to obtain a proper perspective on the manner in which this text-book is likely to influence the training of statisticians and of actuaries in particular. In general, text-books on statistics fall into three main categories: (i) elementary books which treat mainly with descriptive statistics and provide little but general guidance, (ii) handbooks which are designed mainly for practical users who may not be concerned with or trained for the formal development of the statistical tests and techniques, and (iii) advanced texts requiring considerable mathematical ability for their proper understanding. This book clearly aims to provide a course from which the essential principles of modern statistical methods can be properly appreciated, so that the student will be in a position either to use books of class (ii) with a proper understanding of the fundamentals or alternatively to proceed to books of class (iii) without undue difficulty. The book must be said to have achieved this aim and the student who has worked through the text and the excellent examples should have a balanced outlook on the subject. Statistics is essentially a practical subject, and in the reviewer's opinion this book strikes a happy balance between practice and the underlying theory.

A few unimportant misprints have been noticed in the text, but some errors have been noted in the illustrative examples of which the following are important:

On p. 45 the mean of Ex. 12·1 and δ of Ex. 12·2 are incorrect; these are unfortunate errors, as these two examples are used throughout Chapter 12 for illustrating various statistical tests and thus give rise to other errors. On p. 215 the values of m_3 and m_4 should be $-3{\cdot}9076$ and $136{\cdot}0218$ respectively, with consequential minor adjustments to Tables 17·2 and 17·5. On p. 220 in Ex. 17·3 the right-hand side of the first equation should be 3253, which renders the solution of the equations in error.

Various errors have been found among the exercises at the end of each chapter and the following might give rise to some difficulty to a student.

Exercises	Question no.	
11	1	Solution should read $\Gamma(l+m)/\Gamma(l)\,\Gamma(m)$.
11	11	Question should read $(\omega-1)(\omega+2)^2$.
12	12	For 1·66 and 2·80 in solution read ·83 and 1·40.
13	8	For a_{10} in solution read 2·829 in place of 1·629.
15	7	For 1·24 in solution read 1·113.
16	16	For ·3904 in solution read ·435.
16	17	In solution 4 and 264 degrees of freedom should be 7 and 261 respectively, etc.
17	5	For 8·84 in solution read 26·52.
17	6	For 78·48 in solution read 235·44.

However, these and the other blemishes can be corrected in a second edition, which this book should justify.

R. E. BEARD

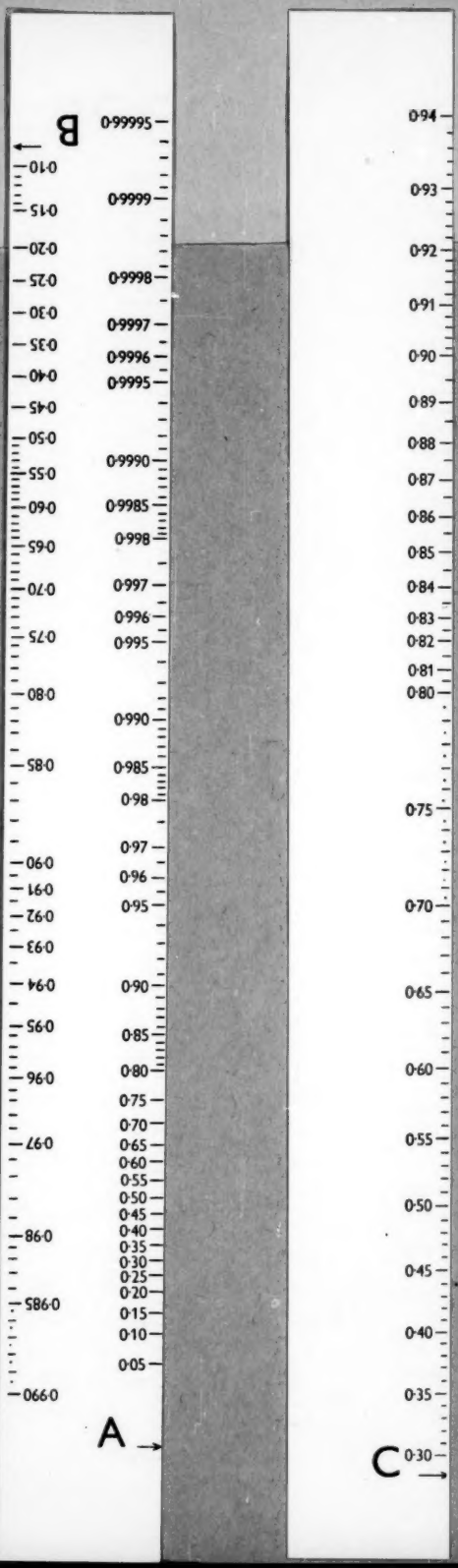
Probability and the Weighing of Evidence. By I. J. Good. vi + 103 pp. Charles Griffin and Co. 1950. Price 16*s*.

The aim of the author has been 'to provide a consistent theory of probability that is mathematically simple, logically sound and adequate as a basis for scientific instruction, for statistics and for ordinary reasoning'. The author maintains that subjective probability judgements must be given a recognized place in the fundamentals of probability and builds his theory on this basis.

At the beginning we have a brief critical review of various theories of probability which have been put forward, including the axiomatic approach which is adopted by the majority of workers in the probability field. This is followed by a discussion of the choice of axioms suitable to act as the basis of a theory, and a mathematical formulation of these axioms, with the attendant theorems. Applications of probability theory to the weighing of evidence and to general statistical theory are discussed at some length.

The corpus of statistical theory has reached such large dimensions during recent years that it is well for statisticians to have reminders, such as this book, of the comparatively insecure foundations on which such an imposing theory is built. It is fair to say that statisticians in general will not accept Dr Good's approach to statistical problems, although those who give Bayes' theorem a prominent part in the theory of estimation may be stimulated by his work.

F. N. DAVID



Two copies of the scale for use with the chart on page 424 are enclosed in the envelope attached.

(*All Rights reserved*)

BIOMETRIKA. Vol. 38, Parts 3 and 4

CONTENTS

PAGE

Biometrika, 1901–1951 267

The estimation of population parameters from data obtained by means of the capture-recapture method. I. The maximum likelihood equations for estimating the death-rate. By P. H. Leslie and Dennis Chitty 269

On estimating the size of mobile populations from recapture data. By Norman T. J. Bailey . . 293

A mathematical theory of animal trapping. By P. A. P. Moran 307

Some systematic experimental designs. By D. R. Cox 312

The comparison of several groups of observations when the ratios of the population variances are unknown. By G. S. James 324

On the comparison of several mean values: an alternative approach. By B. L. Welch 330

The effect of standardization on a χ^2 approximation in factor analysis. By M. S. Bartlett . . 337

The Jacobians of certain matrix transformations useful in multivariate analysis. Based on lectures of P. L. Hsu at the University of North Carolina, 1947. By Walter L. Deemer and Ingram Olkin 345

Two applications of bivariate k-statistics. By M. B. Cook 368

Interrelations between certain linear systematic statistics of samples from any continuous population. By G. P. Sillitto 377

The delay to pedestrians crossing a road. By J. C. Tanner 383

Further applications of range to the analysis of variance. By H. A. David 393

The fitting of polynomials to equidistant data with missing values. By H. O. Hartley 410

Truncated lognormal distributions. I. Solution by moments. By H. R. Thompson 414

A chart for the incomplete beta-function and the cumulative binomial distribution. By H. O. Hartley and E. R. Fitch 423

The expected frequencies in a sample of an animal population in which the abundances of species are log-normally distributed. Part I. By P. M. Grundy 427

Tables of symmetric functions. Parts II and III. By F. N. David and M. G. Kendall 435

Miscellanea

Moment constants for the distribution of range in normal samples. By H. O. Hartley and E. S. Pearson 463

Rapid methods for estimating correlation coefficients. By L. N. Chown and P. A. P. Moran . 464

Mean and variance of an entry in a contingency table. By W. L. Stevens 468

The Wishart distribution derived by solving simultaneous linear differential equations. By H. S. Steyn 470

Quality control systems based on inaccurately measured variables. By W. M. Long 472

The bivariate normal integral. By J. H. Cadwell 475

Note on the variate differences of autoregressive series. By Bento Murteira 479

Note on the inversion theorem. By J. Gil-Pelaez 481

A note on the statistical independence of quadratic forms in the analysis of variance. By J. A. Nelder 482

Reviews

N. L. Johnson and H. Tetley's 'Statistics', vol. II 484

I. J. Good's 'Probability and the Weighing of Evidence' 485

A volume of *Biometrika* containing about 400 pages, with plates and tables, will be published annually in two half-yearly issues.

Papers for publication should either be sent to

PROFESSOR E. S. PEARSON, Department of Statistics, University College, London, W.C. 1,

or if more convenient to

Dr John Wishart, Statistical Laboratory, St Andrew's Hill, Cambridge.
Professor M. G. Kendall, London School of Economics, Houghton St., London, W.C. 2

It is a condition of publication in *Biometrika* that the paper shall not already have been issued elsewhere, and will not be reprinted without leave of the Editors.

Contributors receive 25 copies of their papers free. Joint authors 15 copies each.

The subscription price, *payable in advance*, is Inland 45*s*. *net* per volume and Abroad 54*s*. *net* (including packing and postage). At present certain early volumes are out of print, but where available the cost is Vols. 1 to 32, £5 each, and Vols. 33 to 37, 63*s*. each (excluding postage). Index to Vols. 1 to 5, 2*s*. *net*. Index to Vols. 1 to 15, 5*s*. *net*. Bound volumes, of those volumes in print, are again available at a charge of 12*s*. 6*d*. extra per volume; also binding cases at 7*s*. 6*d*. each. Cheques must be made payable to *Biometrika*, crossed "*a/c Biometrika Trust*" and sent to The Secretary, *Biometrika* Office, Department of Statistics, University College, London, W.C. 1, to whom all orders for series, single copies and offprints should be addressed. All foreign cheques must be drawn in sterling and on a Bank having a London Agency.

Printed in Great Britain at the University Press, Cambridge Brooke Crutchley, University Printer)